COHORT 3 – IMPACT EVALUATION REPORT

The Effects of an Accelerated Learning Program on Out-of-School Children's Academic and Social-Emotional Outcomes in Northeast Nigeria

INTERNATIONAL RESCUE COMMITTEE

Foreign, Commonwealth & Development Office

This research report was prepared by Dr. Jeongmin Lee, Researcher for Education and Dr. Silvia Diazgranados Ferrans, Associate Director for Educational Research at the International Rescue Committee (IRC). The work is the result of a collaborative effort. We would like to thank Yilkal Chalachew, Adane Miheretu, Daniel Sunday, Isaac Olugbenle, Maazu Andrew and Muhammed Bashir for their field work and thoughtful feedback. We would like to also thank learning facilitators, students, and enumerators who participated in data collection. Finally, we would like to thank the United Kingdom Foreign, Commonwealth & Development Office (FCDO) who made this study possible.

Introduction

The Boko Haram insurgency and out-of-school children in Nigeria

Guaranteeing the right to "an inclusive and equitable quality education at all levels" by 2030 is one seminal goal of Sustainable Development Goals (UNESCO, 2015, p. 29). However, only 18% of conflict-affected and fragile countries are on track to meet this goal (Samman et al., 2018). The progress is particularly slow in Nigeria. Since 2009, the Boko Haram insurgency has crippled the country. An estimated 1.9 million people have been displaced and the country's already fragile education system has been devastated: 1,200 schools have been destroyed and 1,700 have closed. More than 600 teachers have been killed and about 19,000 have been displaced (OCHA, 2018).

The number of out-of-school (OOS) children is greater in Nigeria than in any other country in the world. More than 10.5 million children and youth are OOS nationwide. In Nigeria's northeast states such as Yobe and Borno, the damages caused by the recent Boko Haram crisis are most pronounced. In these states 75% of school-aged children have lost access to school (OCHA, 2018). Even when children are in school, their learning is often compromised by a poor school environment as reflected in overcrowded classrooms, poor infrastructure, lack of teaching and learning materials, shortage of qualified teachers and low quality of instruction (FMoE, 2014). Furthermore, traumatic experiences stemming from conflict, poverty and displacement negatively impact children's ability to learn (Shonkoff et al., 2012).

PHOTO CREDIT: THE IRC

PHOTO CREDIT: THE IRC

The outbreak of COVID-19 and its implications for children's education in Nigeria

The COVID-19 pandemic has further complicated the difficulties of ensuring quality education for children in Nigeria. In compliance with the directive of the Federal and State governments for prevention of the spread of COVID-19, all schools and other learning spaces such as non-formal learning centres had to stay closed across the country, including Borno and Yobe States. Since October 2020, schools and learning institutions gradually started to reopen. Still, many educators and students have been hesitant or slow to return to school due to a concern about the ongoing pandemic and poor hygiene and sanitation facilities in school. Essentially, with the outbreak of the COVID-19 pandemic and the chronic problems related to poor quality education, the challenges children in Nigeria have been facing have gotten worse.

IRC's Accelerated Learning Program (ALP) in Northeast Nigeria

ALP curriculum and instructional support activities

While the complexity of the educational challenges in Nigeria requires manifold interventions, research shows that programs that provide safe, accessible non-formal learning spaces with supportive instructors can help children learn and thrive. Building on this evidence base, the International Rescue Committee (IRC) with support from the United Kingdom's Foreign, Commonwealth & Development Office (FCDO) has implemented a nine-month comprehensive accelerating learning program (ALP) for OOS children in Borno and Yobe. Since its inception in 2017, the program has created 400 non-formal learning centres (NFLC) in the two states and served over 34,000 OOS children ages between 9-14. Children eligible for the program had dropped out of school for more than two years or had never been in school. Children enrolled in the ALP are supported to learn the foundational literacy and numeracy skills they need to transition to the formal education system and social and emotional learning (SEL) skills they need to cope with difficult circumstances in life and thrive in life.

The ALP provides classes 3 days per week, 3 hours per day, for 9 months, conducted by trained learning facilitators. Each class has three 45-minute sessions with literacy, numeracy and SEL lessons. Literacy lessons focus on teaching foundational reading skills such as phonological awareness, alphabetic knowledge, fluency, and comprehension. Numeracy lessons include numerical concepts such as magnitude, ranking and counting as well as numerical operations such as addition and subtraction. SEL lessons promote children's ability to manage emotions and positive relationships, such as emotional recognition, empathy, perseverance, stress and impulse management and conflict resolution. The SEL lessons also help children practice executive functioning skills such as focus and memory.

The learning facilitators (LFs) were recruited from their respective community as volunteers because they were proficient in literacy and numeracy. LFs receive a monthly stipend for their work and attend a six-day face-to-face training workshop where they learn pedagogical knowledge in literacy, numeracy and SEL instruction as well as various strategies for lesson planning, classroom management and positive discipline. LFs also attend monthly teaching learning circles (TLC) in which they share useful instructional methods and ideas with their colleagues. TLCs take place monthly and each meeting lasts 2 hours. Each LF has approximately 40 learners in their classroom.

Finally, the ALP staff mobilize support from communities for program implementation and monitoring activities. These members address concerns about the program operations within communities and promote a sense of safety and security, especially among women and girls.

PHOTO CREDIT: THE IRC

Adaptation in response to the program effects observed with the first cohort of the program and the COVID-19 pandemic

An impact evaluation study with the first cohort of the ALP participants in 2018 showed that overall, the program was a cost-effective solution to improving OOS children's literacy and numeracy skill development, though its effect was inconclusive for SEL outcomes. The learning facilitator (LF) coaching component of the ALP did not establish evidence as being supportive of enhanced learning. In a follow-up qualitative study, children and LFs in the APL reported literacy and numeracy being the most enjoyable and practical subjects. This compared to SEL which felt difficult to teach and learn for LFs and learners alike. For LFs, the teaching of the SEL curriculum was less straightforward due to novelty in instructional concepts and methods. For learners, SEL was enjoyable because the subject gave them opportunities to learn and share feelings and experience with friends and LFs. Still, the novelty of the SEL curriculum contributed to lower uptake and lower quality of implementation (IRC, 2019).

The findings from the first cohort of the ALP, with an extended funding from the FCDO, rendered an opportunity for the IRC to review and adjust the previous ALP program and strengthen the quality of LF instruction and support systems. Improvement strategies included:

- **At the student level:** attendance support by providing students with school snacks to incentivize regular attendance.
- **At the LF level:** instruction quality assurance by hiring LFs who have proficiency in Hausa–a medium of instruction in the ALP, providing LFs with teaching manuals in English in addition to Hausa manuals, and adding more emphasis on SEL during pedagogy training and coaching support.
- **At the coaches level:** coaching quality assurance by having coaches attend LF trainings so that coaches learn pedagogical principles promoted in the ALP and reinforce these principles during LF coaching sessions, offering coaches a three-day refresher training on effective coaching strategies as well as ALP curricular principles with an increased emphasis on SEL, and having coaches utilize coaching strategies that are associated with positive changes in LF instruction and student learning, such as engaging LFs in setting coaching goals and encourage LFs' reflection on teaching practices.

In the current report, we initially sought to examine the effects of the revised ALP on OOS children's academic and social-emotional skill development. Yet, as we further detailed below, the implementation of the third cohort of the ALP coincided with the Covid-19 pandemic and the ensuing lockdown and school closure. The implications of these problems for the present study were multifold, ranging from intervention compliance, implementation conditions, data collection, and program evaluation to result interpretation.

Among all, some of the greatest challenges were the inevitable discontinuation of on-site coaching, a temporary break of other program activities and compromised program fidelity. To mitigate these challenges, the IRC Nigeria provided children registered in the program with provisional activities to support their home/community-based learning and safety during the pandemic besides the initially planned ALP activities from April to September in 2020. These support activities were:

PHOTO CREDIT: THE IRC

- **Community-based small group learning circles:** During the six months school closure, the IRC organized 1,683 small learning groups in communities across Borno and Yobe states with the help of 300 LFs. The effort was well received by children and parents. 12,542 (4,977 male: 7,565 female) children attended the learning circles and continued learning during the pandemic. Each learning circle included 5 to 10 children, who took literacy and numeracy lessons three days per week, with the help of LFs. The learning circles observed preventive measures against the spread of COVID-19, including social distancing and handwashing practices. The program distributed 266 handwashing stations and other hand washing supplies such as plastic water containers and soaps for use during the small community-level learning circle sessions. Still, parents' concerns about COVID-19 were significant. However, attendance rates at the learning circles were not as robust as the rate observed at the ALP before the pandemic.
- **Provision of story books:** The 12,542 children were also provided with about 19,000 story books written in Hausa, a language most widely spoken in Borno and Yobe to encourage them to read at home, independently or with support from any family members who could read those books to them such as older siblings or parents. Each child received two books.
- **Provision of 1,400 radio sets:** The 12,542 children were provided with pre-recorded literacy and numeracy lessons that were broadcasted by the State Universal Basic Education Boards (SUBEB) in Borno and Yobe in collaboration with FHI360, an influential international development organization active in Nigeria and UNICEF. Radio sets were distributed in clusters, where a group of 3 to 5 learners shared one radio set. For learners who were out of reach of the radio broadcast, the IRC staff collected the pre-recorded lessons from FHI360 and saved them on memory cards inserted into radio sets and distributed them to the learners.

The present study

The initial purpose of this research report was to examine the effects of the adjusted model of the ALP on OOS children's learning outcomes in literacy, numeracy, and SEL. Yet, with the anticipated outbreak of Covid-19, we expanded this initial goal and incorporated the three provisional learning supports during school closure in our analysis. As such, we address the following four research questions:

- **RQ1:** What are the baseline-endline changes in the proportions of children at different performance levels in literacy, numeracy and SEL outcomes by treatment status?
- **RQ2:** What are the effects of the ALP implemented during Covid-19 on children's literacy, numeracy and SEL outcomes?
- **RQ3:** How do the supplementary learning supports provided to the treatment and comparison groups during a school lockdown moderate the effects of the ALP on children' literacy, numeracy, and SEL outcomes?
- **RQ4:** Do the effects of the ALP implemented during Covid-19 vary for subgroups of students divided by gender, socio-economic status (SES), migration status, disability, and mother tongue language (MTL)?

It is important to note that the program effects reported in this paper necessitate nuanced interpretations given that the ALP was implemented amid the unprecedented pandemic and its implementation fidelity, dosage, and uptake were inevitably compromised as a result. In other words, the effects of the ALP on children's learning observed in this study are confounded by various noises created by the pandemic, despite various analytic strategies we used to attenuate their impacts. This makes it difficult to precisely tell how the ALP would have performed in the absence of such a challenge. Nevertheless, the findings of this study will still be instrumental for policymakers and practitioners in Nigeria and beyond in understanding how an educational program such as the ALP analysed in this study may support OOS children's learning in the time of Covid-19 and what additional supports should be accompanied to optimize benefits.

PHOTO CREDIT: THE IRC

Methods

The present study used a randomized comparisonled trial to identify the impact of the ALP program implemented by the IRC for OOS children in Borno and Yobe states in Northeast Nigeria. The treatment group attended the ALP for 9 months in total with 7 months prior to the outbreak of the Covid-19 pandemic and the ensuing school closure in April 2020 and 2 months after the lift of school closure in October in 2020. The comparison group was placed in a waitlist to attend the program at a later time. During the school closure, the three provisional learning supports the IRC provided were available to both groups. Table 1 presents the treatment contrast between treatment and comparison groups.

Table 1. Treatment contrast

Initial Intervention activities	Treatment	Comparison
Three-day face-to-face trainer of trainers training	✓	—
Four-day face-to-face facilitator training and two-day refresher	✓	—
Teaching learning circles, once a month for two hours	✓	—
One-day community coalition training and one-day refresher	✓	—
Civic society organization face-to-face training	✓	—
Learning materials—notebooks, bags, pens, erasers	✓	—
Instructional materials for learning facilitator—teaching guides and lesson plans for literacy, numeracy and SEL	✓	—
Classroom equipment—teaching aids, blackboard, chalks, work-books, cleaning tools, and attendance book	✓	—
Classes for children—three-hour sessions per day, three times per week for nine months	✓	—
Stipend for facilitators	✓	—
***Added learning support during school closure due to Covid-19**		
Community learning circles 3 times per week	✓	✓
Provision of 1,400 radio sets and radio lessons 3 times per week	✓	✓
Provision of 1,900 story books	✓	✓

PHOTO CREDIT: THE IRC

Participants

The target population of the ALP is OOS children between the ages nine to sixteen who missed the chance to complete primary education and have stayed out of school for two years or longer in Borno and Yobe states in Northeast Nigeria. Since its inception the ALP has served the total of 33,883 OOS children in 400 NFLCs in the two states. During the third cohort, which is the focus of this report, the program provided services to 14,097 children, of which 13,528 (8,354 female: 5,174 male) graduated in December 2020.

The research sample of this study was drawn from the 13,528 children in the third cohort. To recruit children for the research, researchers randomly selected 43 centres from the 400 NFLCs. Given oversubscription at each NFLC, some children eligible to receive services were randomly assigned to attend the ALP courses (treatment) and others were placed in a wait list (comparison). It is important to note that although children in the comparison group did not attend the ALP courses during this study, the program did not preclude them from seeking alternative educational opportunities. Some children in the waitlist group may have attended a formal school, a non-formal education program implemented by other development agencies, or a religion-based education institute. Therefore, in this study the causal effects of the ALP show the learning gains children were able to achieve by attending the ALP courses relative to the gains they may have achieved in a natural setting, not a strict laboratory experiment.

Within this random treatment assignment structure, the ALP team surveyed a total of 1,327 children at baseline in November 2019 (684 children from the treatment group and 644 from the comparison group). However, after 7 months into the program, the ALP had to close due to the COVID-19 outbreak in early 2020 and remained closed until the state governments in Borno and Yobe lifted an order on school closure in October 2020. During closures, all children both in treatment and comparison groups received the provisional learning activities described above (learning circles, radio lessons, and storybooks to read at home). The ALP activities resumed in September 2020.

At endline in February 2021, about 2 months after program resumption, we re-assessed learning outcomes among 885 of the 1,327 children included in the baseline sample, which led to a 33.3 attrition rate across the whole sample. Among the final sample, 466 were in the treatment group (31.7% attrition) and 419 were in the comparison group (35.4% attrition). This sample size was sufficient to detect a minimum effect size of 0.23 standard deviations (SD) with 80% power at the 5% significance level. The average age of participants was 12 years. 43% were male; 57% were female. About 72% spoke Hausa language at home; 28% spoke Kanuri. 42% were displaced children, while 58% were from host communities. Table 2 presents the number of children and NFLC for the treatment and comparison groups.

Table 2. The number of children and learning centres involved in the present study

	Baseline (n = 1,327)		Endline (n = 885)	
	Treatment	**Comparison**	**Treatment**	**Comparison**
Students	683	644	466	419
NFLCs	43	(43)	41	(41)

** The number of NFLCs in the comparison was the same as that in the treatment.*

PHOTO CREDIT: THE IRC

To ensure the treatment and comparison groups were balanced in initial learning outcomes and background characteristics that could affect treatment outcomes such as age, gender, families' SES levels, language spoken at home among others, we conducted a series of t-tests with the baseline sample after accounting for children's clustering within each centre. Overall, we found the two groups statistically differed in two of nine background characteristics and two of 21 learning outcomes with the differences. These group differences in baseline characteristics became more evident after attrition, leading to additional differences in children's mother tongue language (MTL) and several numeracy outcomes including number discrimination, addition level 1, and subtraction level 1. These differences were consistently in favor of the treatment group over the comparison group. Specifically, compared to the comparison group, children in the treatment group were older by 0.3 years (about 4 months). They were more likely to report speaking Hausa at home and having been to school before by 2 percent, respectively. They also had higher scores in letter sound identification, number discrimination, addition level 1, and subtraction level 1. The magnitudes of these baseline differences were small but statistically significant. This pattern could have inflated the effects of ALP. To comparison for potential biases, we included covariates in regression models to estimate treatment effects. Appendix 1 shows the balance statistics at baseline.

Instruments

Background information questionnaire: This instrument collected children's demographic information, such as sex, SES, IDP and disability status.

Early Grade Reading Assessment (EGRA): Literacy skills were assessed through five subtasks, including timed letter name identification, timed non-word reading, timed oral reading fluency, listening comprehension (Reliability coefficient = 0.80) and reading comprehension (Reliability coefficient = 0.83).

Early Grade Mathematics Assessment (EGMA): Numeracy skills were assessed through eight subtasks including timed number identification, number discrimination (Reliability coefficient = 0.91), missing number (Reliability coefficient = 0.86), timed addition level 1, timed subtraction level 1, addition level 2 (Reliability coefficient = 0.84), subtraction level 2 (Reliability coefficient = 0.82), and word problems.

Battery of social and emotional learning (SEL) tools: SEL skills were assessed using a battery of assessments that captured children's hostile attribution bias (Reliability coefficient = 0.64), two types of emotional intensity i.e., anger (Reliability coefficient = 0.77) and sadness (Reliability coefficient = 0.79), emotional attribution accuracy (Reliability coefficient = 0.55), depression (Reliability coefficient = 0.88), and three types of intensity toward conflict resolution, i.e., aggression (Reliability coefficient = 0.84), disengagement (Reliability coefficient = 0.86) and problem solving (Reliability coefficient = 0.88).

Procedure

The data were collected by a team of trained local enumerators who had worked with the IRC and other development organizations in Northeast Nigeria in various education projects. These enumerators had ample experience in child assessments. All of them received five days of training before baseline data collection in November 2019. In the training, the enumerators learned the logistics of the project, the ethical principles they should follow during child assessment, and rehearsed how to administer assessments in tablets. The endline data was collected in February 2021. The enumerators attended a five-day refresher training prior to the field work and were reminded of the data collection strategies they learned at the initial training. The study was reviewed and received IRB approval by INGO's Research Ethics Committee and the National Health Research Ethics Committee of Nigeria (NHREC).

Analytic approach

RQ1: To examine baseline-endline changes in the proportions of learners at different performance levels, we used a percentage statistic and charted (a) the percentages of zero-score learners—those who scored zero scores in each literacy, numeracy and SEL outcome, (b) the percentages of below-target-score learners—those who scored below the target scores determined in this study, and 3) the percentages of above-target-score learners—those who scored above the target scores determined in this study at baseline and endline.

Literacy target scores are 70 for letter sound and non-word reading, 60 for oral reading fluency, and 80 for listening comprehension and reading comprehension. Numeracy target scores are 20 for addition level 1 and subtraction level 1, 30 for number identification, and 80 for number discrimination, missing number, addition level 2, subtraction level 2, and word problems.

In SEL, sadness and anger intensity scores are high for scores ≥ 2, medium for $1 \leq$ scores < 2 and low for $0 \leq$ scores < 1. Hostile attribution, disengagement, problem-solving, aggression and emotional attribution accuracy scores are high for scores ≥ 0.67, medium for $.33 \leq$ scores $< .67$ and low for $0 \leq$ scores $< .33$. Depression scores are high for scores ≥ 12, medium for $1 \leq$ scores < 12 and low for score $= 0$.

RQ2: To determine the effects of the ALP on key outcomes of interest, we used a difference-in-difference (DID) regression model, after controlling for student covariates including age, gender, physical disability, SES, and home literacy environments as well as the clustering of students within each learning centre.

RQ3: To estimate the moderation effects of the three supplementary learning supports on key outcomes of interest, we ran the DID regression model fitted in RQ2 after adding an interaction term between a variable denoting treatment effect and a dummy variable denoting the uptake of each learning support (yes/no).

RQ4: To estimate differential treatment effects of the ALP on key outcomes of interest for different subgroups of children, we ran the DID regression model fitted in RQ2 and added an interaction term between a variable denoting treatment effect and a dummy variable denoting a subgroup category (female/male, low/high SES, IDP/host, disabled/able-bodied, and Hausa/Kanuri)

Results

RQ1: Baseline-endline changes in the proportions of learners by performance levels

Literacy

Changes in the percentages of students with zero scores: The baseline data showed that about 50-70 percent of children in the treatment and comparison groups obtained zero scores in all literacy outcomes except in listening comprehension, where only five to six percent had zero scores. At endline, the percentages of zero-score learners decreased in all literacy outcomes in both groups, with the greatest decrease observed in letter sound identification (a 40.3 percent point decrease in the treatment group: a 32.7 percent point decrease in the comparison group). A cross-group comparison showed that overall, the magnitudes of decrease were larger in the treatment group than in the comparison group.

Changes in the percentages of students who met target scores: At baseline, the percentages of children who scored higher than the project 's target scores ranged about 12-13 across all literacy outcomes in both the treatment and comparison groups except listening comprehension where the proportions of these children was about 72 percent. At endline, the percentages of above-target-score learners increased in listening comprehension (a 12.4 percent point increase in the treatment group: a 4.8 percent point increase in the comparison group) and reading comprehension (a 24.0 percent point increase in the treatment group: a 2.4 percent point increase in the comparison group). The percentages decreased in other outcomes such as non-word reading (a 12.9 percent point decrease in the treatment group: an 18.7 percent point decrease in the comparison group). A cross-group comparison showed that overall, the magnitudes of increase were larger in the treatment group than in the comparison group while the magnitudes of decrease were smaller in the former group than in the latter.

A concurrent decrease in the proportions of children meeting the learning standards in both groups may be explained by learning losses associated to the COVID-19 pandemic. Studies from high-income countries such as the United States repeatedly show (e.g., Atteberry & McEachinthat, 2020) children lose up to 40 percent of the gains they have made over the school year while on a multi-month vacation. This summer break or vacation effect ultimately means that learning loss reverses some of the learning achievement students have made over the year. The effect of the school break, in this case caused by the COVID-19, seemed also pronounced in the present study. As such, many children who met the learning at baseline suffered a learning loss over a seven-month school closure. The proportion of these children moderately decreased at endline across all groups. Figure 1 below shows the baseline-endline changes in the proportions of children at different performance levels in two key literacy outcomes, i.e., oral reading fluency and reading comprehension, by group.

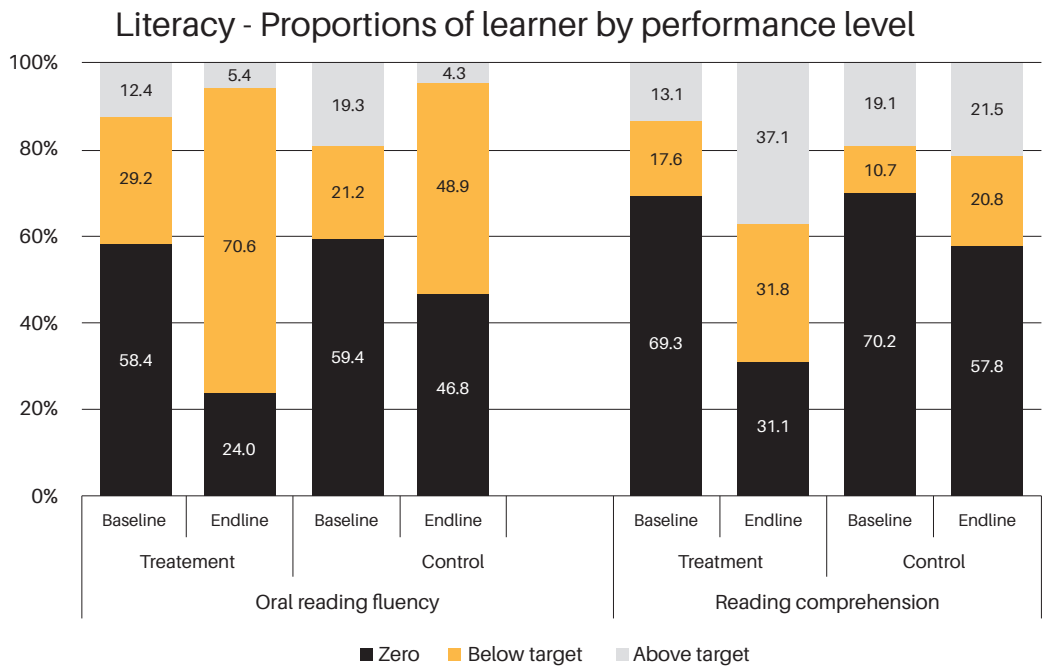


Figure 1. Proportions (%) of learners at different performance levels in oral reading fluency and reading comprehension at baseline and endline by treatment assignment

Numeracy

Changes in the percentages of students who obtained zero scores: At baseline, the percentages of children with zero scores varied across different numeracy outcomes with the greatest percentage observed in subtraction level 2 (about 54-55 percent) and the smallest percentage observed in word problems (about 11-13 percent) in both the treatment and comparison groups. At endline, the percentages of zero-score learners decreased in all numeracy outcomes in both groups. The magnitudes of decrease were different for different outcomes for different groups. Within these variations, the greatest decrease in zero scores was observed in addition level 2 (a 38.4 percent point decrease) and the smallest decrease was observed in word problems (an 8.3 percent point decrease) both in the treatment group.

Changes in the percentages of students who met target scores: At baseline, the percentages of children who scored higher than the project 's target scores were different for different numeracy outcomes. The greatest percentage observed was in word problems (about 55-57 percent) and the smallest percentages were observed in addition level 1 and subtraction level 1 (about 12-20 percent) in both groups. At endline, the percentages of above-target-score learners increased in some outcomes such as number discrimination (a 39.3 percent point increase in the treatment group: a 14.5 percent point increase in the comparison group) and addition level 2 (a 78.5 percent point increase in the treatment group: a 1.9 percent point increase in the comparison group).In other numeracy outcomes, the pattern was reversed likely due to a learning loss over the seven-month school closure. The percentages decreased in addition level 1 (a 2.6 percent point decrease in the treatment group: a 13.9 percent point decrease in the comparison group) and subtraction level 1 (a 4.3 percent point decrease in the treatment group: a 16.5 percent point decrease in the comparison group) in both groups. A cross-group comparison showed that overall, the scope and magnitudes of increase were larger in the treatment group than in the comparison group. Figure 2 presented below shows the baseline-endline changes in the proportions of children at different performance levels in addition level 1, subtraction 1 and word problems by group.

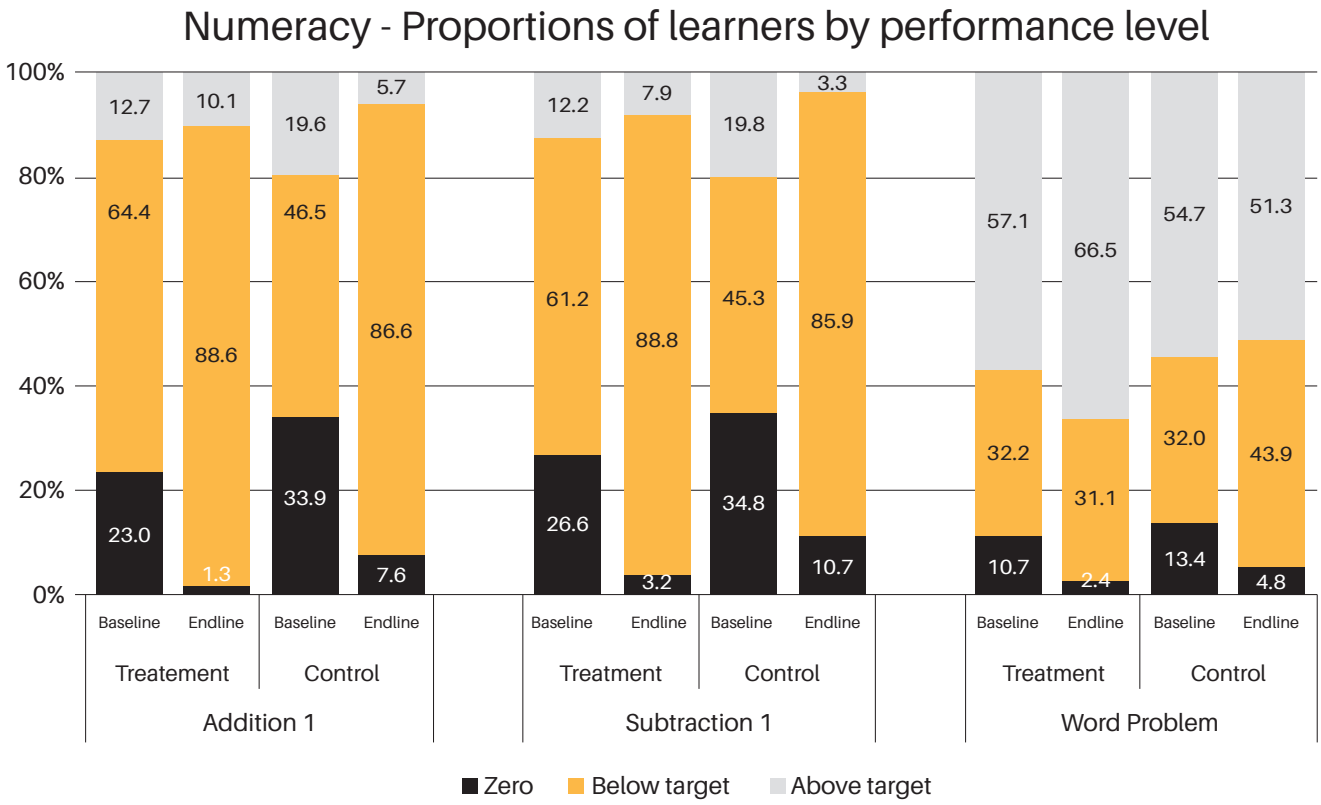


Figure 2. Proportions (%) of learners at different performance levels in addition level 1, subtraction level 2, and word problems at baseline and endline by treatment assignment

SEL

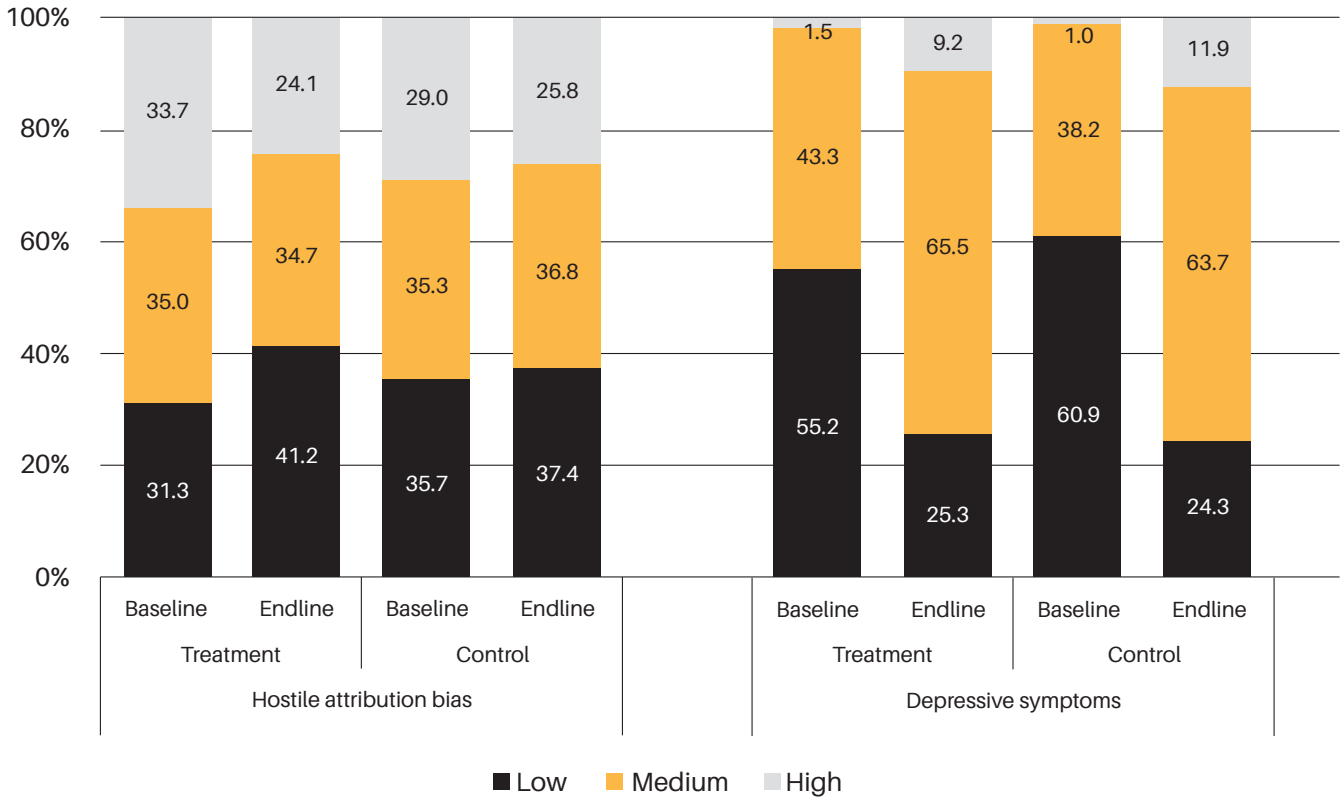
In this section, the SEL outcomes analysed in this brief were grouped into (a) socially desirable outcomes such as emotional attribution accuracy and problem-solving and (b) socially less desirable outcomes such as hostile attribution bias, sadness intensity, anger intensity, aggression, or depressive symptoms. The first paragraph to follow describes the baseline-endline changes in the percentages of children at a lower score level in socially desirable outcomes by group. The second paragraph describes the baseline-endline changes in the percentages of children at a higher score level in socially less desirable outcomes by group. This categorization was meant to examine the upward change in the number of children who were at a low spectrum of socially desirable SEL skills and the downward change in the number of children who were at a high spectrum of socially less desirable SEL skills, by treatment status.

Changes in the percentages of students at a lower-score level for emotional attribution accuracy and problem-solving: When all else is equal, higher scores would be desired for emotional accuracy and problem-solving. At baseline, we observed about 81-84 percent of children in both the treatment and comparison groups were at a lower score level in problem-solving. In emotional attribution accuracy, no child was at a low score level and they all were at a medium- or high-score level. At endline, these percentages decreased in problem-solving in both groups while they remained nearly the same in emotional attribution accuracy. Figure 3 below shows the baseline-endline changes in the proportions of children at different score levels in two of the socially less desirable SEL outcomes, i.e., hostile attribution bias and depressive symptoms, by group. We included more detailed information about the percentages of children at different learning levels by time and treatment status in Appendix 2.

Changes in the percentages of students at a higher-score level for hostile attribution bias, sadness intensity, anger intensity, and depression: When all else is equal, lower scores would be desired for hostile attribution bias, sadness intensity, anger intensity, and depression. At baseline, the percentages of children who were a higher score level in these items varied, with the greatest percentage observed in hostile attribution bias (about 30-34 percent) and the smallest percentage observed in depressive symptoms (about 1-2 percent) in both the treatment and comparison groups.

At endline, these percentages decreased in most of these outcomes including hostile attribution bias (a 8.6 percent point decrease in the treatment group: a 2.7 percent point decrease in the comparison group), sadness intensity (a 7.5 percent point decrease in the treatment group: a 7.8 percent point decrease in the comparison group), and anger intensity (a 3.7 percent point decrease in the treatment group: a 0.2 percent point decrease in the comparison group), although they increased in depressive symptoms (a 7.7 percent point increase in the treatment group: a 10.9 percent point increase in the comparison group).

Figure 3. Proportions (%) of learners at different performance levels in hostile attribution bias and depressive symptoms at baseline and endline by treatment assignment



- Based on the observations reported above, we conclude from baseline to endline that coincided with the COVID-19 pandemic:
- **Literacy:** The proportions of children who obtained zero scores decreased in the treatment and comparison groups alike. However, the proportions of children meeting project targets increased only in some outcomes such as listening comprehension and reading comprehension in both groups.
- **Numeracy:** The proportions of children who obtained zero scores decreased in both the treatment and comparison groups. However, the proportions of children meeting project targets increased more in the treatment group than the comparison group. These outcomes included number discrimination, number discrimination, missing number, subtraction level 2, addition level 2, and word problems.
- **SEL:** For pro-social SEL outcomes such as emotional attribution accuracy and problem- solving, higher scores would be desirable. But, we observed the proportions of children who were at a high-score level slightly decreased in problem-solving from baseline to endline in both the treatment and comparison groups. The proportions remained nearly the same in emotional attribution accuracy in both groups.
- For outcomes that reflect internalizing and externalizing problems, such as hostile attribution bias and depressive symptoms, lower scores would be desirable. We found mixed patterns. For hostile attribution bias, the proportion of children at a lower-score level increased in both group while those at a high score level decreased. However, for depressive symptoms, the pattern is reversed.

RQ2: Average treatment effects of the ALP on children's learning outcomes

In literacy, the ALP had statistically significant and positive effects on nearly all outcomes except listening comprehension. At endline, children in the treatment group, compared to those in the comparison group, scored higher in letter sound identification by 7 points (per minute), non-word reading by 6 points (per minute), oral reading fluency by 8 points (per minute), and reading comprehension by 18 percent points. Effect sizes varied with the largest effect size observed in reading comprehension (0.43 SD) and the smallest effect size observed in letter sound identification (0.26 SD).

In numeracy, the ALP had statistically non-significant effects on all outcomes except number identification. In this outcome, the treatment group scored higher than the comparison group by 4 points (per minute) leading to the effect size of 0.29 SD.

Lastly, in SEL the ALP had statistically non-significant effects on all SEL outcomes. Figure 4 visually presents the effect sizes of ALP for different learning outcomes. We marked statistically non-significant treatment effects with translucent bars to contrast against statistically significant effects. Appendix 3 shows more detailed statistics about group averages across time and program effects.

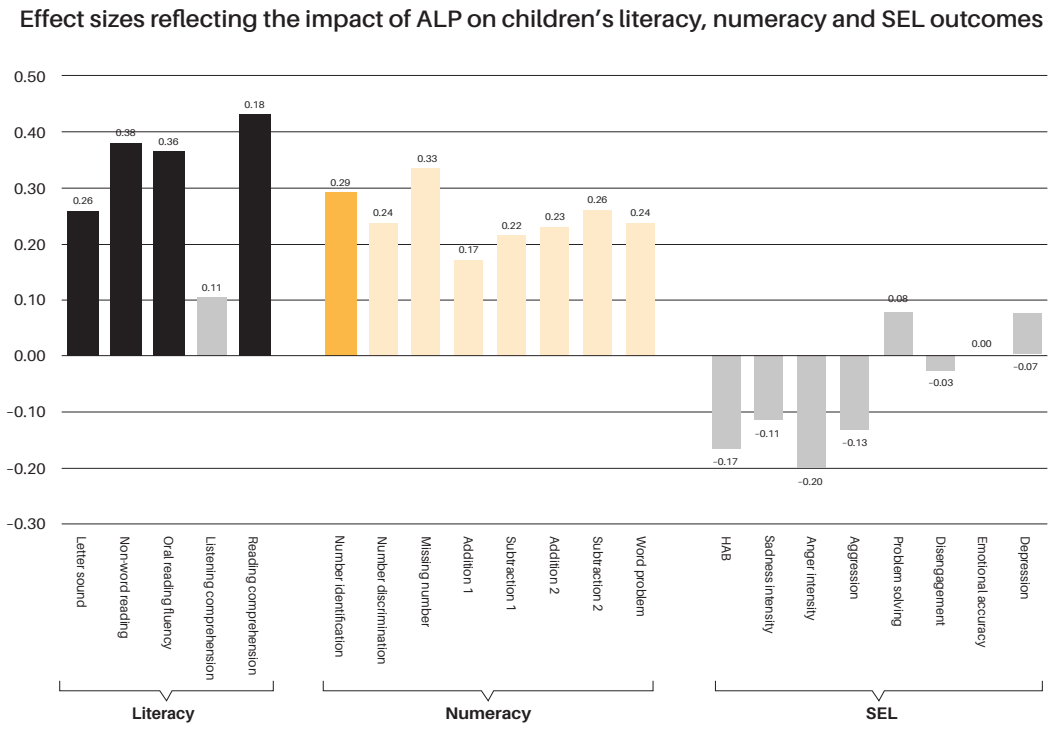


Figure 4. Effect sizes reflecting the impact of the ALP disrupted by Covid on children' literacy, numeracy and SEL outcomes. *Translucent bars indicate that treatment effects were statistically non-significant.

Based on the findings above, we conclude that the ALP effectively improved children's literacy skills despite various implementation challenges caused by the COVID-19 pandemic. However, such improvement was not achieved in numeracy or SEL outcomes except number identification.

RQ3: Moderation effects of the supplementary learning supports on treatment effects

Overall, the availability of supplementary learning supports during school closure—(a) small learning circle, (b) radio lesson, (c) independent reading of storybooks, and (d) family's support for reading storybooks—did not alter the effects of the ALP on children's learning outcomes. In other words, how the ALP affected children's learning outcomes was not reliant on whether they received the supplementary learning supports or not during the school closure in general. Still, having access to small group learning circles beside the initial ALP courses helped children gain higher scores in the number discrimination item by 22 percent points. It also helped lessen anger intensity by 58 percent points. These statistics mean that when children had continued learning support during the school closure via small group learning circles, they were able to benefit from the ALP in the two outcome areas when the program resumed. We included more detailed statistics about the moderation effects of the supplementary learning supports on treatment effects in Appendix 4.

Based on the above findings, we conclude that despite sizable disruptions caused by the COVID-19 pandemic, when children had continued learning support through small group learning circles, they were able to reap greater benefits from the ALP, thought mostly in number discrimination and anger regulation.

RQ4: Differential treatment effects for different subgroups

The baseline data showed that some subgroups had greater disadvantages over other groups across literacy, numeracy and SEL outcomes. At endline, we observed differential treatment effects among some subgroups across literacy, numeracy and SEL outcomes as well. We detail these findings for each subgroup below.

By sex: At baseline, female students underperformed male students in several literacy and numeracy outcomes, such as oral reading fluency, number identification, additional level 2, and word problems. At endline, we noticed female students benefited from the ALP more than male students did in several outcomes across literacy, numeracy and SEL, including reading comprehension, number discrimination, additional level 2, and anger regulation, although male students benefitted more in terms of amicably solving interpersonal conflict. When the baseline differences and the different treatment effects were put together, we noticed that the ALP addressed the baseline learning gap between female and male students in the addition level 1 item in numeracy.

By SES: At baseline, children in the lower SES group underperformed children in the higher SES group in several literacy, numeracy and SEL outcomes, including listening comprehension, word problems, aggressive reaction to interpersonal conflict, and emotional accuracy. At endline, we observed that children in the higher SES group benefited more from the ALP than children in the lower SES group did in a number of literacy, numeracy, and SEL outcomes. These outcomes included oral reading fluency, reading comprehension, number identification, discrimination, missing number, addition level 2, subtraction level 2 and aggression. The pattern was reversed in the anger regulation item in SEL, however. When the baseline differences and the different treatment effects combined, we noticed that the baseline learning gap between the lower and higher SES groups in aggression were widened rather than reduced.

PHOTO CREDIT: THE IRC

PHOTO CREDIT: THE IRC

By IDP: At baseline, IDP children academically outperformed children from host communities in nearly all literacy and numeracy outcomes except the listening comprehension item in literacy and the word problems item in numeracy. They, however, scored lower in the emotional attribution accuracy item in SEL. At endline, there was no statistically different treatment effects by displacement status in literacy or numeracy. In SEL, however, the treatment effects were greater among IDP children than among their counterparts in the disengagement and depression items. However, these results especially for the depression item necessitated a nuance interpretation. As we described in sections 5.1-5.2, the average scores of depression and the proportion of children at a high level of depression increased among the whole research sample. Given this pattern, positive and larger treatment effects for IDP children mean that IDP children may have experienced severer depression during the pandemic than children from host communities despite SEL support provided through the ALP.

By disability: At baseline, there were no statistically significant group differences in literacy or numeracy outcomes by disability status. Yet, in the children with a physical disability scored higher than able-bodied children in the depression item in SEL. At endline, disabled children benefitted more from the ALP than able-bodied children did in various numeracy outcomes, including number discrimination, missing number, addition level 1, subtraction level 1, addition level 2 and subtraction level 2. When the baseline differences and the different treatment effects combined, we noticed that the ALP had no impact on addressing the baseline disparity in depression by disability status.

By MTL: At baseline, children who speak Kanuri language at home compared to those who speak Hausa scored higher in the number discrimination item in numeracy and the hostile attribution bias in SEL. At endline, we did not observe statistically different treatment effects in any of the learning outcomes, whether literacy, numeracy, or SEL. When the baseline and endline findings combined, we noticed that the ALP had no impact on addressing the baseline learning differences by MTL. In Table 3, we marked subgroups who were benefited more from the ALP and changes in the baseline equity learning gaps. We included more detailed statistics about the baseline differences and the differential treatment effects by subgroup in Appendix 5.

Table 3. Groups who benefited more from the ALP, and changes in the baseline learning gaps

Outcome		ALP vs Comparison									
		Sex		SES		IDP		Disability		MTL	
		Girl	Boy	Lower	Higher	IDP	Host	Disabled	Able-bodied	Non-Hausa	Hausa
EGRA	Letter identification						▼				
	Non-word reading						▼				
	Oral reading fluency	▼			▲		▼				
	Listening comprehension			▼							
	Reading comprehension	▲			▲		▼				
EGMA	Number identification	▼			▲		▼				
	Number discrimination	▲			▲		▼	▲			▼
	Missing number				▲		▼	▲			
	Addition 1						▼	▲			
	Subtraction 1						▼	▲			
	Addition 2	▲	▼		▲		▼	▲			
	Subtraction 2				▲		▼	▲			
	Word problems			▼							
SEL	Hostile attribution bias									▼	
	Sadness intensity										
	Anger intensity			▲							
	Aggression			▼	▲						
	Problem solving		▲								
	Disengagement						▲				
	Emotional attribution accuracy			▼							
	Depressionw						▲	▼			

Note: ▼ indicates groups who were more disadvantaged at baseline. ▲ indicates groups who benefited more from the ALP. Yellow cell indicates groups where we observed reductions in baseline equity gaps. Grey cell indicates groups where we observed the reinforcement of baseline equity gaps.

The effects of the ALP varied by children's gender, SES levels, displacement status and MTL with benefits often in favor of female and disabled children from higher SES families in host communities. Within these variations, the ALP reduced the baseline difference between female and male students in the additional level 2 item in numeracy. But the difference in average aggression scores was widened by family's SES background as children from higher SES families showed a greater reduction in this area than those from lower SES families did.

PHOTO CREDIT: THE IRC

Conclusions

ALPs have been used in a variety of educational settings from the global North to the global South as a solution to providing over-aged OOS children with access to quality learning opportunities. However, information is lacking if this approach could be also effective in emergency settings. The present research study sought to fill this knowledge gap by examining the effects of an ALP implemented by the IRC in Borno and Yobe states in Nigeria.

The impact evaluation with the first cohort of the ALP showed that the program was a cost-effective solution to improving OOS children's literacy and numeracy skill development, though its effect was inconclusive for SEL outcomes. Based on these findings the IRC reviewed and adjusted the initial ALP model and strengthened the quality of LF instruction and support systems with an extended funding from the FCDO. This adjusted model of the ALP was implemented with the third cohort of the program from September 2020 to December 2020, despite the disruption of the regular ALP activities between April and October 2020, due to COVID19 pandemic. The current study examined the effects of this adjusted ALP via an RCT study.

Results showed that the adjusted model of ALP was effective in improving OOS children's literacy skills in various areas encompassing letter sound recognition, non-word reading, oral reading fluency and reading comprehension. In terms of numeracy skills, the program also effectively improved children's ability to accurately identify numbers. These findings were consistent with a prior RCT study conducted with the first cohort of the ALP in 2018 where positive program effects were documented in various literacy and numeracy outcomes.

However, the ALP did not lead to any statistically significant improvement in children's SEL outcomes. There could be various reasons for this. Among all, the COVID-19 pandemic may have been critical. This speculation was partly supported by substantial score increases in some of SEL outcomes such as depression, sadness intensity, and anger intensity during the pandemic across the whole research sample regardless treatment assignment. This pattern indicates that it might be difficult for the ALP to reverse the negative effects of the learning losses and psychosocial distress caused by the pandemic on children's SEL outcomes, even if the program could have been effective in normal circumstances.

In addition, in analysing the data from an equity lens, we also found that the effects of the ALP varied by children's gender, SES levels, displacement, and disability status. The benefits often were in favor of female and disabled children from higher SES families in host communities. Some of these differential treatment effects contributed to reducing the baseline learning gap between female and male students in the addition level 2 item in numeracy, although the baseline gap was intensified in the aggression item in SEL.

Limitations

There are a few limitations in this study to acknowledge. The first limitation is related to the internal validity of the study in terms that despite randomization of children to treatment, we observed several differences in background characteristics and initial learning levels between the treatment and comparison groups. In addition, we had about 24% of attrition at endline. This attrition exacerbated the group imbalance observed previously, causing another threat to internal validity. Both issues could introduce a bias in the estimation of program effects. To minimize these threats, we used the DID analysis. This approach was helpful to separate treatment effect from the change that might have occurred in the absence of treatment, in this case represented by the comparison group. We also included influential background covariates in our regression model. Despite these two strategies, it is still possible that factors that were unobserved or omitted in our regression might drive the results.

Another internal validity threat was related to history, which refers to external factors or events that coincide in time with the intervention and may masquerade as treatment effect (Shadish et al., 2002). In our study, the COVID-19 pandemic and the supplementary learning supports provided during school closure act as confounders. However, the two events were not local to only one group. Rather, both groups experienced the pandemic and had an equal access to the learning supports. Therefore, we do not anticipate history has a vast and uneven effect. Also, we proactively incorporated the supplementary learning supports during school closure in our analysis and examined their influences on intervention outcomes as shown earlier.

However, treatment strength was still severely compromised by the COVID-19 pandemic. The situation was further complicated by relevant challenges such as a long break amid intervention implementation as well as weakened students' and LF's attendance and retention. Essentially, these factors may have weakened treatment strength and deteriorated implementation conditions.

Finally, it is important to note that the causal relationships reported in this study may not be generalized to ages, settings, or regions beyond the population represented in our study's research sample. As such, the findings are limitedly generalizable to out-of-school children between age 9-16 in Yobe and Borno states in Nigeria. In addition, the implementation of ALP analysed in this study coincided with the COVID-19 pandemic. Thus, the effects reported here may not reflect what would be observed in other circumstances. In fact, the effects of ALP tested among the first cohort of children in 2018 were much larger across literacy, numeracy, and SEL outcomes (Ferrans et al, 2019). As such, the program effectively improved children's scores in letter identification (ES = 0.29 SD), non-words reading (ES = 0.26 SD), oral reading fluency (ES = 0.23 SD), number identification (ES = 0.41 SD), number discrimination (ES = 0.43 SD), missing number (ES = 0.58 SD), addition level 1 (ES = 0.29 SD), subtraction level 1 (ES = 0.32 SD), addition level 2 (ES = 0.46 SD), subtraction level 2 (ES = 0.23), and word problems (ES = 0.37 SD). Also, it helped reduce orientation toward disengaging from conflicts (ES = -0.18 SD).

Photo caption. You can add photos to your document, but please do not add borders or drop shadows or any other embellisments.

Recommendations and next steps

Providing equitable access to quality education to all children is a vital goal of SDGs (UNESCO, 2015). However, progress has been slow toward achieving this goal in conflict-affected settings. The challenge is particularly pronounced in Northeast Nigeria, where Boko-Haram insurgencies have devastated an already weak government education system and 75% of school-aged children have lost access to school as a result (OCHA, 2018).

Adequate funding must be directed to providing these children with access to school, through combinations of government funds and international aid as needed. While this is necessary, we would like to also emphasize that the process may take a number of years. In this situation, the findings of this study offer evidence that a non-formal ALP program that combines literacy, numeracy and SEL instructions is one effective alternative approach. We urge policymakers and donors to consider more robust and large-scale investment in this program for OOS children in Northeast Nigeria and potentially in other regions when the formal system is hard to reach.

Also, the study's findings related SEL highlight the need for more substantial funding for comprehensive mental health programs to mitigate the negative psychosocial effects of numerous challenges faced by children in emergency settings such as poverty, orphanhood, and violence. In sections 5.1 and 5.2 as well as in Appendix 3, we reported the implementation of the present study coincided with the outbreak of the COVID-19 pandemic. The scores of several SEL outcomes such as anger intensity, sadness intensity and depression increased despite the provision of the ALP.

This pattern is indicative of exacerbated mental health during the recent pandemic. At the same time, it suggests that the teaching of adaptive SEL strategies included in the initial ALP alone was limited to prevent children from experiencing such a problem. On the other hand, despite a multi-month school closure, when children had provisional learning support at home and in the community, they were able to demonstrate improvement in a few numeracy and SEL outcomes. It may be worthwhile to provide such support in psychosocial well-being as well.

References

Atteberry, A., & McEachin, A. (2020). School's out: The role of summers in understanding achievement disparities. (EdWorkingPaper: 19-82). Retrieved from https://doi.org/10.26300/2mam-bp02

Ferrans, S.D., Lee, J., Reilly, A., & Miheretu, A. (2019). The cost-effectiveness of ALP for out-of-school children and the added value and effect of providing on-site coaching to learning facilitators in NFLC on students' learning and transition outcomes in Nigeria: Endline report. International Rescue Committee.

FMoE. (2014). Nigeria: Education for all 2015 national review. Federal Ministry of Education. http://planipolis.iiep.unesco.org/en/2014/nigeria-efa-review-report-2000-2014-strides-and-milestones-6007

OCHA. (2018). Humanitarian response plan—Nigeria. Office for the Coordination of Humanitarian Affairs.

Samman, E., Lucci, P., Hagen-Zanker, J., Bhatkal, T., Simunovic, A. T., Nicolai, S., Stuart, E., & Caron, C. (2018). SDG progress: Fragility, crisis and leaving no one behind. Overseas Development Institute.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). Experimental and quasi-experimental designs for generalized causal inference. Houghton, Mifflin, and Company.

Shonkoff, J., Garner, A., Committee on Psychosocial Aspects of Child and Family Health, Committee on Early Childhood, Adoption, and Dependent Care, & Section on Developmental and Behavioral Pediatrics. (2012). The lifelong effects of early childhood adversity and toxic stress. Pediatrics, 129(1), 232–246.

UNESCO. (2015). World Education Forum 2015: Final report. UNESCO.

Appendix 1. Baseline-endline changes in the proportions (%) of children at different performance levels

Variable	Sample mean (SD) (n = 885)	Treatment (n = 446)	Comparison (n = 419)	Treat-Comparison
Background characteristics				
Age	11.96 (1.95)	11.15	10.74	0.41*
Female	0.57 (0.50)	0.59	0.54	0.05
Disabled	0.07 (0.25)	0.06	0.08	-0.02
Displacement status	0.42 (0.49)	0.46	0.38	0.07
Hausa language spoken at home	0.72 (0.45)	0.68	0.77	-0.09*
School repetition experience	0.34 (0.47)	0.27	0.42	-0.16*
Living in a rental house	0.57 (0.49)	0.56	0.55	0.01
Low socioeconomic status	0.53 (0.50)	0.52	0.52	0.00
Low home learning environment	0.66 (0.47)	0.85	0.90	-0.05
Literacy				
Letter identification	34.93 (28.29)	10.58	5.86	4.73*
Non-words reading	13.96 (15.19)	3.90	2.29	1.61
Oral reading fluency	21.47 (22.45)	6.16	3.46	2.70
Listening comprehension	0.85 (0.25)	0.79	0.77	0.02
Reading comprehension	0.39 (0.41)	0.08	0.05	0.03
Numeracy				
Number identification	19.76 (12.81)	10.22	7.85	2.36
Number discrimination	0.74 (0.29)	0.45	0.35	0.10*
Missing number	0.48 (0.30)	0.24	0.19	0.06
Addition L1	11.77 (6.04)	6.21	4.56	1.65*
Subtraction L1	10.95 (5.61)	5.57	4.04	1.53*
Addition L2	0.51 (0.35)	0.24	0.18	0.06
Subtraction L2	0.44 (0.34)	0.20	0.16	0.04
Word problems	0.74 (0.29)	0.66	0.62	0.04
SEL				
Hostile attribution bias	0.45 (0.30)	0.50	0.47	0.03
Sadness intensity	1.20 (1.15)	0.94	0.79	0.15
Anger intensity	0.84 (1.14)	0.65	0.57	0.08
Disengagement	0.04 (0.15)	0.28	0.26	0.02
Problem-solving	0.34 (0.38)	0.54	0.56	-0.01
Aggression	0.62 (0.39)	0.17	0.18	-0.01
Emotional accuracy	0.92 (0.12)	0.83	0.83	0.00
Depression	4.00 (4.05)	1.53	1.42	0.11

Note: Statistical significance is marked as *p<.05 **p<.01 *** p<.001.

Appendix 2. Baseline-endline changes in the proportions (%) of children at different performance levels

Outcome	Level	Baseline		Endline		Changes	
		Treatment	Comparison	Treatment	Comparison	Treatment	Comparison
LITERACY							
Letter sound	Zero	47.0	53.7	6.7	21.0	-40.3	-32.7
	Below target	41.2	28.2	79.0	69.5	37.8	41.3
	Above target	11.8	18.1	14.4	9.5	2.6	-8.6
Non-word reading	Zero	59.2	59.4	24.2	47.5	-35.0	-11.9
	Below target	27.7	21.7	75.5	52.3	47.8	30.6
	Above target	13.1	18.9	0.2	0.2	-12.9	-18.7
Oral reading fluency	Zero	58.4	59.4	24.0	46.8	-34.4	-12.6
	Below target	29.2	21.2	70.6	48.9	41.4	27.7
	Above target	12.4	19.3	5.4	4.3	-7.0	-15.0
Listening comprehension	Zero	4.7	6.0	3.0	3.8	-1.7	-2.2
	Below target	23.8	21.7	13.1	19.1	-10.7	-2.6
	Above target	71.5	72.3	83.9	77.1	12.4	4.8
Reading comprehension	Zero	69.3	70.2	31.1	57.8	-38.2	-12.4
	Below target	17.6	10.7	31.8	20.8	14.2	10.1
	Above target	13.1	19.1	37.1	21.5	24.0	2.4
NUMERACY							
Number identification	Zero	17.0	24.8	1.3	6.9	-15.7	-17.9
	Below target	69.5	54.7	72.3	79.3	2.8	24.5
	Above target	13.5	20.5	26.4	13.8	12.9	-6.7
Number discrimination	Zero	19.7	32.5	1.9	4.5	-17.8	-28.0
	Below target	47.2	33.7	25.8	47.0	-21.4	13.3
	Above target	33.0	33.9	72.3	48.4	39.3	14.5
Missing number	Zero	24.2	37.9	5.4	9.8	-18.8	-28.1
	Below target	61.8	41.5	67.0	75.7	5.2	34.2
	Above target	13.9	20.5	27.7	14.6	13.8	-5.9
Addition 1	Zero	23.0	33.9	1.3	7.6	-21.7	-26.3
	Below target	64.4	46.5	88.6	86.6	24.2	40.1
	Above target	12.7	19.6	10.1	5.7	-2.6	-13.9
Subtraction 1	Zero	26.6	34.8	3.2	10.7	-23.4	-24.1
	Below target	61.2	45.3	88.8	85.9	27.6	40.6
	Above target	12.2	19.8	7.9	3.3	-4.3	-16.5
Addition 1	Zero	48.9	51.3	10.5	22.0	-38.4	-29.3
	Below target	28.8	24.1	48.7	51.6	19.9	27.5
	Above target	22.3	24.6	40.8	26.5	18.5	1.9

Subtraction 2	Zero	54.9	53.9	16.7	32.9	-38.2	-21.0
	Below target	24.7	23.4	49.8	46.1	25.1	22.7
	Above target	20.4	22.7	33.5	21.0	13.1	-1.7
Word problems	Zero	10.7	13.4	2.4	4.8	-8.3	-8.6
	Below target	32.2	32.0	31.1	43.9	-1.1	11.9
	Above target	57.1	54.7	66.5	51.3	9.4	-3.4
SEL							
HAB	Low	31.3	37.0	39.9	34.1	8.6	-2.9
	Medium	34.8	32.9	34.8	38.4	0.0	5.5
	High	33.9	30.1	25.3	27.4	-8.6	-2.7
Sadness intensity	Low	30.0	36.5	27.9	24.8	-2.1	-11.7
	Medium	24.7	22.0	22.7	27.9	-2.0	5.9
	High	31.1	30.5	23.6	22.7	-7.5	-7.8
	Other feelings	14.2	11.0	25.8	24.6	11.6	13.6
Anger intensity	Low	50.4	54.4	54.7	50.6	4.3	-3.8
	Medium	25.1	25.1	17.6	15.3	-7.5	-9.8
	High	14.6	11.9	10.9	11.7	-3.7	-0.2
	Other feelings	9.9	8.6	16.7	22.4	6.8	13.8
Aggression	Low	69.7	71.6	96.4	95.7	26.7	24.1
	Medium	5.2	6.4	1.3	1.9	-3.9	-4.5
	High	25.1	22.0	2.4	2.4	-22.7	-19.6
Disengagement	Low	39.7	38.4	32.8	28.6	-6.9	-9.8
	Medium	8.2	10.3	10.9	14.3	2.7	4.0
	High	52.1	51.3	56.2	57.0	4.1	5.7
Problem-solving	Low	83.9	80.9	60.9	62.8	-23.0	-18.1
	Medium	4.1	5.5	10.3	12.2	6.2	6.7
	High	12.0	13.6	28.8	25.1	16.8	11.5
Emotional accuracy	Low	0.0	0.0	0.0	0.2	0.0	0.2
	Medium	10.9	16.0	3.2	3.6	-7.7	-12.4
	High	89.1	84.0	96.8	96.2	7.7	12.2
Depressive symptoms	Low	55.2	60.9	25.3	24.3	-29.9	-36.6
	Medium	43.3	38.2	65.5	63.7	22.2	25.5
	High	1.5	1.0	9.2	11.9	7.7	10.9

Note: Literacy target scores are 70 for letter sound and non-word reading, 60 for oral reading fluency, and 80 for listening and reading comprehension. Numeracy target scores are 20 for addition 1 and subtraction 1, 30 for number identification, and 80 for number discrimination, missing number, addition 2, subtraction 2, and word problems. In SEL, sadness and anger intensity scores are high for scores ≥ 2, medium for 1 ≤ scores < 2 and low for 0 ≤ scores < 1. Hostile attribution, disengagement, problem-solving, aggression and emotional attribution accuracy scores are high for scores ≥ 0.67, medium for .33 ≤ scores < .67 and low for 0 ≤ scores < .33. Depression scores are high for scores ≥ 12, medium for 1 ≤ scores < 12 and low for score = 0.

Appendix 3. Average treatment effects

	Baseline			Endline			Changes	
	Treatment	Comparison	Diff	Treatment	Comparison	Diff	Treatment	Comparison
LITERACY								
Letter sound (per min)	11.20	6.43	4.77	40.14	28.06	12.08	7.31* (3.48)	0.26
Non-word reading (per min)	3.65	3.07	0.59	16.78	10.43	6.35	5.76* (1.80)	0.38
Oral reading fluency (per min)	5.65	4.50	1.15	25.71	16.41	9.31	8.15* (3.17)	0.36
Listening comprehension (%)	0.79	0.78	0.01	0.86	0.82	0.04	0.03* (0.05)	0.11
Reading comprehension (%)	0.07	0.07	-0.01	0.47	0.30	0.17	0.18** (0.06)	0.43
NUMERACY								
Number identification (per min)	9.89	8.38	1.51	22.19	16.96	5.23	3.72* (1.82)	0.29
Number discrimination	0.44	0.37	0.07	0.81	0.67	0.14	0.07* (0.05)	0.24
Missing number	0.24	0.20	0.04	0.54	0.40	0.14	0.10* (0.06)	0.33
Addition 1 (per min)	6.23	4.66	1.57	12.97	10.35	2.62	1.04* (0.98)	0.17
Subtraction 1 (per min)	5.57	4.25	1.32	12.07	9.54	2.53	1.21* (1.05)	0.22
Addition 2 (%)	0.24	0.19	0.05	0.57	0.44	0.13	0.08* (0.06)	0.23
Subtraction 2 (%)	0.20	0.16	0.04	0.50	0.37	0.13	0.09* (0.07)	0.26
Word problem (%)	0.67	0.62	0.04	0.79	0.68	0.11	0.07* (3.48)	0.24
SEL								
HAB (avg)	0.51	0.47	0.03	0.44	0.46	-0.02	-0.05 (0.04)	-0.17
Sadness intensity (avg)	0.96	0.76	0.20	1.23	1.16	0.07	-0.13 (0.14)	-0.11
Anger intensity (avg)	0.65	0.57	0.09	0.77	0.91	-0.14	-0.23 (0.16)	-0.20
Aggression (avg)	0.27	0.27	0.00	0.03	0.05	-0.02	-0.02 (0.04)	-0.13
Problem solving (avg)	0.55	0.55	0.01	0.36	0.32	0.04	0.03 (0.06)	0.08
Disengagement (avg)	0.17	0.18	-0.01	0.61	0.63	-0.02	-0.01 (0.05)	-0.03
Emotional attribution accuracy (avg)	0.83	0.83	-0.01	0.92	0.92	-0.01	0.00 (0.02)	0.00
Depression (total)	1.49	1.49	0.00	3.85	4.14	-0.29	-0.29 (0.32)	-0.07

Note: Standard errors are in parentheses. Statistical significance is marked as *p<.05 **p<.01 *** p<.001.

Appendix 4. Moderation effects of learning supports during a lockdown on treatment effects

Outcome	Moderator				
	Small learning circle (yes/ no)	Radio lesson (yes/no)	Independent reading (yes/no)	Reading support (yes/no)	Overall suport level* (high/low)
LITERACY					
Letter sound (per min)	5.27 (4.53)	-5.69 (5.49)	-7.89 (5.52)	-2.86 (4.65)	-0.03 (4.44)
Non-word reading (per min)	-1.21 (2.59)	-0.94 (2.54)	0.58 (3.05)	0.15 (2.40)	1.36 (2.35)
Oral reading fluency (per min)	1.57 (4.25)	-2.26 (3.94)	-6.00 (5.38)	1.16 (3.99)	1.21 (4.38)
Listening comprehension (%)	0.01 (0.08)	0.02 (0.06)	-0.03 (0.10)	0.06 (0.06)	0.02 (0.06)
Reading comprehension (%)	0.01 (0.08)	-0.02 (0.08)	-0.15 (0.10)	0.04 (0.08)	0.03 (0.09)
NUMERACY					
Number identification (per min)	3.96 (2.85)	-2.20 (2.67)	-3.96 (2.81)	-0.35 (2.33)	-1.06 (2.32)
Number discrimination	0.22* (0.10)	-0.13 (0.09)	-0.17 (0.10)	-0.09 (0.08)	-0.07 (0.09)
Missing number	0.07 (0.08)	-0.09 (0.08)	-0.15 (0.10)	-0.02 (0.07)	0.06 (0.07)
Addition 1 (per min)	0.67 (1.72)	-1.19 (1.78)	-1.43 (1.78)	-0.48 (1.52)	1.24 (1.56)
Subtraction 1 (per min)	0.10 (0.11)	-0.14 (0.09)	-0.21 (0.12)	-0.00 (0.09)	0.06 (0.09)
Addition 2 (%)	0.09 (0.11)	-0.14 (0.10)	-0.14 (0.12)	-0.01 (0.08)	0.06 (0.09)
Subtraction 2 (%)	0.10 (0.11)	-0.14 (0.09)	-0.21 (0.12)	0.00 (0.09)	0.06 (0.09)
Word problem (%)	0.01 (0.11)	-0.03 (0.10)	-0.08 (0.10)	0.03 (0.07)	0.05 (0.08)
SEL					
HAB (avg)	-0.13 (0.08)	0.16 (0.09)	0.12 (0.08)	0.17 (0.07)	0.12 (0.09)
Sadness intensity (avg)	0.06 (0.27)	-0.14 (0.29)	-0.11 (0.28)	0.18 (0.25)	0.06 (0.24)
Anger intensity (avg)	-0.58* (0.25)	0.66 (0.33)	0.56 (0.29)	0.33 (0.29)	0.48 (0.30)
Aggression (avg)	-0.03 (0.07)	0.03 (0.09)	0.04 (0.08)	-0.02 (0.06)	0.04 (0.06)
Problem solving (avg)	0.05 (0.10)	-0.07 (0.10)	-0.08 (0.13)	-0.01 (0.09)	0.04 (0.08)
Disengagement (avg)	-0.02 (0.08)	0.04 (0.09)	0.04 (0.10)	0.03 (0.07)	-0.08 (0.06)
Emotional attribution accuracy (avg)	-0.01 (0.04)	-0.04 (0.03)	-0.01 (0.04)	-0.00 (0.03)	-0.00 (0.03)
Depression (total)	1.03 (0.80)	-0.56 (0.81)	1.04 (0.71)	0.05 (0.81)	-0.33 (0.87)

Note: a. Overall support level aggregates the sum score of the four types of learning support provided during school lockdown. The higher-level support group indicates children who reported the update of three or more supports and the lower-level support group indicates children who reported the uptake of two or less supports. Standard errors are in parentheses. Statistical significance is marked as *p<.05 **p<.01 *** p<.001.

Appendix 5. Baseline differences and differential treatment effects by subgroup

	Baseline group mean differenes					Differential treatment effects				
	Female	Lower SES	IDP	Disabled	Kanuri	Female	Lower SES	IDP	Disabled	Kanuri
LITERACY										
Letter sound (per min)	-1.90 (0.98)	-0.49 (2.18)	5.02* (2.11)	-0.64 (2.13)	-2.08 (2.47)	6.12 (3.87)	-18.13 (10.00)	-5.42 (4.09)	9.99 (5.32)	-0.93 (5.48)
Non-word reading (per min)	-0.80 (0.52)	0.10 (0.96)	2.46** (0.88)	0.19 (1.03)	-0.92 (1.07)	3.05 (1.84)	-4.05 (3.76)	-2.04 (2.23)	5.86 (2.97)	1.04 (2.60)
Oral reading fluency (per min)	-2.62* (1.19)	1.31 (1.64)	4.07* (1.76)	0.04 (1.29)	-2.11 (1.93)	6.64 (3.57)	-15.53* (7.15)	-2.41 (4.07)	9.66 (4.99)	-0.07 (4.32)
Listening comprehension (%)	-0.01 (0.02)	-0.11** (0.03)	0.02 (0.05)	-0.01 (0.06)	-0.07 (0.05)	0.01 (0.07)	-0.17 (0.10)	-0.08 (0.07)	0.12 (0.11)	-0.03 (0.09)
Reading comprehension (%)	-0.02 (0.01)	0.01 (0.02)	0.06* (0.02)	0.01 (0.02)	-0.02 (0.02)	0.15** (0.05)	-0.26* (0.12)	-0.04 (0.06)	0.19 (0.10)	0.05 (0.08)
NUMERACY										
Number identification (per min)	-2.35** (0.84)	1.35 (1.18)	3.74** (1.23)	-0.42 (1.14)	-0.05 (1.29)	2.49 (1.66)	-9.69** (3.47)	-2.54 (2.44)	4.98 (3.13)	1.38 (3.13)
Number discrimination	-0.04 (0.03)	-0.01 (0.05)	0.11* (0.05)	-0.02 (0.06)	0.18** (0.06)	0.10* (0.05)	-0.24* (0.11)	0.01 (0.10)	0.31* (0.13)	-0.06 (0.10)
Missing number	-0.03 (0.02)	0.01 (0.03)	0.09** (0.03)	0.00 (0.04)	0.02 (0.04)	0.11 (0.06)	-0.20* (0.09)	-0.01 (0.06)	0.27** (0.09)	0.00 (0.07)
Addition 1 (per min)	-0.37 (0.47)	0.16 (0.80)	1.40* (0.68)	0.53 (1.18)	1.51 (0.86)	1.88 (0.95)	-2.01 (1.72)	0.21 (1.38)	4.79* (2.12)	0.43 (1.55)
Subtraction 1 (per min)	-0.37 (0.47)	0.45 (0.69)	1.42* (0.65)	0.05 (0.86)	1.30 (0.83)	1.41 (1.01)	-2.17 (1.64)	0.60 (1.24)	5.96** (1.92)	-0.06 (1.51)
Addition 2 (%)	-0.06* (0.03)	0.01 (0.05)	0.13** (0.04)	-0.01 (0.06)	0.01 (0.05)	0.12* (0.06)	-0.23* (0.10)	-0.06 (0.07)	0.30** (0.11)	-0.03 (0.09)
Subtraction 2 (%)	-0.05 (0.03)	0.03 (0.04)	0.11* (0.04)	0.04 (0.05)	-0.02 (0.04)	0.09 (0.06)	-0.33* (0.13)	-0.04 (0.06)	0.32** (0.11)	-0.08 (0.09)
Word problem (%)	-0.06* (0.03)	-0.17*** (0.04)	-0.01 (0.05)	-0.07 (0.05)	-0.05 (0.07)	0.02 (0.07)	-0.16 (0.12)	-0.07 (0.08)	0.24 (0.12)	-0.05 (0.10)
SEL										
HAB (avg)	0.00 (0.02)	-0.03 (0.05)	-0.07 (0.04)	0.06 (0.05)	0.11* (0.05)	-0.07 (0.05)	-0.11 (0.15)	0.04 (0.07)	0.01 (0.07)	0.03 (0.07)
Sadness intensity	0.00 (0.07)	-0.04 (0.09)	0.04 (0.09)	-0.02 (0.13)	-0.27 (0.15)	0.00 (0.17)	0.08 (0.41)	-0.13 (0.21)	0.23 (0.26)	0.31 (0.24)
Anger intensity	-0.05 (0.06)	0.13 (0.12)	-0.15 (0.09)	-0.05 (0.15)	0.30 (0.16)	-0.35* (0.16)	-1.07** (0.33)	-0.19 (0.20)	-0.45 (0.32)	-0.52 (0.26)
Aggression	-0.02 (0.03)	0.10* (0.05)	0.01 (0.05)	0.05 (0.08)	0.01 (0.06)	0.10 (0.06)	0.29* (0.11)	0.00 (0.07)	0.00 (0.08)	-0.00 (0.08)
Problem solving	-0.01 (0.04)	-0.09 (0.06)	-0.00 (0.06)	-0.07 (0.08)	-0.06 (0.08)	-0.19* (0.08)	-0.15 (0.15)	-0.18 (0.10)	0.01 (0.11)	-0.02 (0.11)
Disengagement	0.03 (0.02)	-0.01 (0.03)	-0.01 (0.03)	0.02 (0.05)	0.05 (0.04)	0.09 (0.06)	-0.14 (0.17)	0.18* (0.07)	-0.01 (0.08)	0.02 (0.08)
Emotional attribution accuracy	0.01 (0.01)	-0.07** (0.02)	-0.08* (0.03)	-0.03 (0.06)	0.03 (0.03)	-0.05 (0.02)	-0.00 (0.07)	0.00 (0.03)	0.06 (0.04)	-0.07 (0.04)
Depression	0.17 (0.14)	0.27 (0.38)	0.41 (0.32)	3.79** (1.25)	0.82 (0.59)	0.92 (0.56)	0.75 (1.41)	1.55** (0.55)	0.69 (0.65)	0.82 (0.79)

Note: Female students vs male students. Lower SES children vs higher SEL children. IDP children vs children from host communities. Disable children vs able-bodied children. Non-Hausa languages speakers vs Hausa language speakers. Standard errors are in parentheses. Statistical significance is marked as *p<.05 **p<.01 *** p<.001.