Understanding Complementary Basic Education in Ghana

Cycle 4 Endline Report

Department for International Development

October 2017

In association with:

Prepared by Jonathan Stern and Jennifer Pressley, with support from the research and management team at the University of Sussex, University of Cambridge, and IMC Worldwide.

Table of Contents

List of Figures

List of Tables

Abbreviations

CBE	Complementary Basic Education
DFID	Department for International Development
DRIC	Directorate of Research Innovation and Consultancy
EGMA	Early Grade Mathematics Assessment
EGRA	Early Grade Reading Assessment
GES	Ghana Education Service
GILLBT	Ghana Institute of Linguistics, Literacy and Bible Translation
IP	implementing partner
IRT	item response theory
JEAVCO	JEAVCO Associates Ltd
LCD	Link Community Development
MoE	Ministry of Education
OLS	ordinary least squares
OOSC	out-of-school children
P1–P6	primary 1 through primary 6 (grades)
PAB	PAB Consult Limited
PCA	principal component analysis
PISA	Program for International Student Assessment
QA	Quality Assurance
RTI	RTI International
TIMSS	Trends in International Mathematics and Science Study
UCC	University of Cape Coast
USAID	United States Agency for International Development

Executive Summary

The Complementary Basic Education (CBE) programme in Ghana is designed to provide 'second-chance' access to basic education for out-of-school children (OOSC) of primary school age. The present study is part of a larger research agenda to estimate the impact of the CBE programme on learning outcomes and access to primary and junior secondary formal education. As such, this report details the results of the endline learner assessment undertaken in June 2017, at the end of the 2016/2017 cycle (i.e. Cycle 4) of the CBE programme.

The methodology used for this study revolved around tracking a sample of students from the start of the CBE cycle in Octoberto the end of the cycle in June 2017. Testing started in early November. Despite the best efforts by the data collection team, it was ultimately not possible to test all 2,401 baseline students at endline, due mostly to dropout from the CBE programme. In the end, 399 students from the baseline sample were unavailable for testing at endline, reducing the total sample to 2,002 students (17% sample attrition). These students were administered the same literacy and numeracy assessments used at baseline, in order to ensure comparability across the two time points. Accordingly, the endline study was designed with two main purposes in mind:

1. Assess the literacy and numeracy gains resulting from the CBE programme
2. Compare the learning gains of students by sex, language, wealth, implementing partner (IP), and region

To maintain consistency with the Cycle 4 baseline report, both the literacy and numeracy assessments were used to create component proficiency scores for students, across five different categories: Basic Reading, Advanced Reading, Writing, Basic Numeracy, and Advanced Numeracy. These scores were scaled from 0 to 100 for ease of interpretation and then divided into the following four proficiency levels:

1. **Non-performer,** comprising those who scored zero on a component score;
2. **Beginner,** comprising those who scored greater than zero but less than 50;
3. **Approaching proficiency,** comprising those who scored greater than 50 but less than 80; and
4. **Proficient,** comprising those who scored greater than 80.

Overall, approximately 80–85% of students in the CBE programme improved in both literacy and numeracy over the nine-month programme. Additionally, there was a significant increase in the mean performance scores for students across all 15 of the subtasks in the literacy and numeracy assessments. While students continued to underperform on the literacy subtasks (as compared with numeracy), the average gains in literacy were higher. On average, mean literacy scores increased by a nearly consistent 25 percentage points. The average increase in numeracy scores showed more variation, with gains ranging from 17 to 29 percentage points.

Results also showed large improvements across all proficiency score categories from baseline to endline. This was apparent in the reductions in the proportions of pupils in the lower-performing categories (non-performer and beginner) and increases in the proportions in the upper-performing categories (approaching proficiency and proficient). More specifically, while 12.6% of students were non-performers in literacy at baseline, by endline that number was cut by two-thirds, to 4.2%. The beginner category also saw a large reduction, from approximately 70% at baseline to approximately 45% at endline. The approaching proficiency category more than doubled, from 12.8% to 27.1%. Perhaps the most impressive finding is that the proficient category more than quadrupled, from 4.8% at baseline to nearly a quarter (24%) at endline. In effect, over 51% of learners were approaching proficiency.

Similar results were found for numeracy. There were reductions in the proportions of non-performers (6% to 2%), and the percentage of beginners was cut in half (from 60% to 30%). While there was only a modest increase in those approaching proficiency (from 26% to 34%), the largest increase was in the proficient category, where the proportion of students rose from 9% at baseline to 35% at endline. These results all point to strong learning gains from the programme. Thus, overall about 69% were proficient or achieving proficiency.

Male students, on average, outperformed female students on the endline assessment. This was due in large part to the greater proportion of male students moving into the proficient category. In other words, while both male and female students improved, it was male students at the higher-performing end of the spectrum who improved most – thus creating a separation in performance by sex in both literacy and numeracy.

No significant differences were found in endline performance or gains based on wealth. However, differences were found in terms of prior school attendance. While students with prior school attendance outperformed their unschooled counterparts at both baseline and endline, the CBE programme was able to narrow this gap. The key finding in terms of gains by language, implementing partner and region, was that relative performance was not consistent over time. In other words, some of the poorest-performing subgroups at baseline actually showed the greatest improvements, while some of the highest-performing groups at baseline had only modest gains. It is difficult from these data to determine what drove these differences (e.g. specific practices, materials), but this is an area that should be investigated further, in order to learn from the languages and implementing partners that had greater gains, and to improve upon those with smaller gains.

In order to determine the impact of various factors on gains in student performance, linear regression models were created to predict improvement across the five competency categories. These models showed that age was a consistently positive predictor of improvement on all outcomes (with older students realising larger gains than younger ones, holding all else constant). Additionally, facilitators helping students when they had a difficult time understanding and engaging in group work were both found to significantly increase gains across a number of competencies. This provides evidence that two of the main programme components (engaged facilitators providing constructive feedback and participating in productive group work activities) are both working to improve learning and should remain at the heart of the programme. Lastly, in terms of language match, these models provided evidence that having a match between a student's language and the centre's language of instruction is particularly important for the most basic skills (i.e. basic reading and basic numeracy)—though further investigation is needed into what can be done to improve learning when language match is low.

In terms of specific languages of instruction, students in Twi schools scored higher on both literacy and numeracy subtasks than students in other languages at baseline, but Brifo and Kasem students showed significantly larger gain scores than their Twi counterparts on all five competencies. Gonja, Dagaare, Kusaal, and Gurune students also all improved more than Twi students for at least three of the competencies. Sissala and Dagbani were the only two languages whose students had significantly smaller gains than Twi students in any of the models.

Ultimately, it is clear that students in CBE centres significantly improved in their literacy and numeracy scores during their nine-month programme. However, this improvement was not consistent by subgroup. Boys improved more than girls, those without prior school experience improved more than those who had previously attended, and there was a great amount of variation in the relative gains made by students in different languages (and in centres run by different implementing partners). While it appears that these students had become more prepared for the formal school system than they were nine months earlier, comparing their performance to that of traditional public school students in the next round of this study will shed further light on the ultimate impact of the CBE programme on school readiness.

Introduction

The Complementary Basic Education (CBE) programme in Ghana is designed to provide 'second-chance' access to basic education for out-of-school children (OOSC) between the ages of 8 and 14. The programme offers literacy and numeracy instruction in the students' mother tongue in community-based classes of a maximum of 25 participants for nine months. Community volunteers, and in some cases, National Service Personnel, serve as the instructors. The students who graduate from the CBE programme after the nine months are later enrolled into mainstream formal education, typically in either in primary 3 or 4 (P3 or P4). CBE is funded by the Department for International Development (DFID) and the United States Agency for International Development (USAID) and implemented in partnership with the Ministry of Education (MoE) and Ghana Education Service (GES) through Crown Agents, Associates for Change and Education Development Trust. The programme has been in operation since 2012 and has reached over 170,000 students and their families.

This present study is part of a larger research agenda to estimate the impact of the CBE programme on learning outcomes and access to primary and junior secondary formal education. The programme's stakeholders are keen to know how effective the programme is in achieving its goals to provide a sustainable route for OOSC to continue their formal education. Data on the learning outcomes of the CBE graduates, their transition patterns into formal education, their long-term retention and completion rates in formal education, and differences in outcomes, if any, across beneficiary groups, will serve to inform the stakeholders how the CBE programme model is working and how its impact can be maximised in the most cost-efficient way.

As one part in this larger research agenda, a study was undertaken to assess the baseline literacy and numeracy proficiency of the students participating in Cycle 4 of the CBE programme in the 2016/2017 academic year. The current study examined the endline performance of those same students at the completion of the nine-month programme. Specifically, this endline study sought to:

1. assess the change in literacy and numeracy levels of students from the time of enrolment to the time of completion in CBE classes;
2. compare the learning gains of the students by sex, age, socioeconomic background, geographical regions, languages, and implementing partners[1]; and

Accordingly, a child background questionnaire, literacy assessment, and numeracy assessment were administered to a proportionally representative sample of 2,002 students at the end of Cycle 4 of the CBE programme in June 2016. In November 2017, another representative sample of students drawn from Cycle 5 will be assessed and then re-assessed in June 2018, representing the endline assessment for that cycle.

The results of the endline study are reported here. The report is organised as follows:

- Chapter 1 describes the specific questionnaire and assessment instruments used in the June 2017 endline, as well as training and data collection procedures.
- Chapter 2 describes the endline sample (including the sampling procedure and sample attrition).
- Chapter 3 presents the results of the endline assessment overall and by sub-group.
- Chapter 4 examines the relative impact of various factors on the variation in scores.
- Chapter 5 concludes with recommendations for subsequent rounds of data collection.

[1] Note that facilitators, implementation models and implementation partners cannot be examined separately based on the indicators available in these data.

1. The CBE Cycle 4 Endline Assessment

The original assessment instruments used for this study were first developed by the Education Development Trust (formerly CFBT) and further revised by the Directorate of Research Innovation and Consultancy (DRIC) of the University of Cape Coast (UCC).[2] DRIC obtained the English and translated versions of the literacy and numeracy instruments of the previous assessment (2014/2015) from the CBE Management Unit. In addition, instructional materials being used to implement the programme in the various local languages were also obtained to help ensure that the items addressed the appropriate content and construct levels of the programme.

DRIC reviewed the English and translated versions of the literacy and numeracy instruments of the previous assessments. The review ensured that the assessment items meaningfully reflected what literacy and numeracy competencies the learners are expected to attain in the programme. DRIC also held consultations with the National Assessment Unit under the Ghana Education Service (GES) in order to ensure agreement on changes being proposed to the original instrument. This process led to the identification of priority areas for testing in relation to literacy/numeracy before the instrument was finalised. For quality assurance purposes, the translation of the various assessment items into the different mother tongues was done following a test and item specification provided to translators by DRIC. Furthermore, to ensure that the translations were of good quality, specialists teaching each of the 12 languages at the University of Education, Winneba, at Ajumako campus were recruited to do the translations.

After careful consideration, it was decided to use the previously piloted and validated baseline literacy and numeracy instruments for this endline evaluation. This was done to ensure direct comparability across time points, without requiring advanced statistical approaches for test equating (which are difficult and often imprecise with short assessments such as these). Because administrators had complete control over the assessments (and they were not shared directly with IPs or teachers), there was no concern about test leakage. Furthermore, with nine months between test administrations, there was no concern about students memorising the test items and remembering them at the endline.

Finally, as noted in the baseline report, while the instruments used do follow many of the basic principles of the Early Grade Reading and Mathematics Assessments (EGRA and EGMA), one important difference is that no fluency scores were recorded for this assessment. In future rounds it would be possible to include fluency measures but they would not be able to be compared back to Cycle 4 results. A child background questionnaire was added (at baseline only; see next section), to allow for a more contextual and statistically robust analysis of the data.

1.1 Child background questionnaire

Since the same students were assessed at both baseline and endline, the child questionnaire was administered only during the baseline data collection. The child background questionnaire used at baseline (see ***Annex 1***) collected information on the students related to their demographics, family status, household economic situation, school and work history, and personal opinions about school and learning. It was designed to permit the analysis of patterns of differences in performance linked to the students' background.

1.2 Literacy assessment instrument

The literacy assessment instrument (see sample in ***Annex 2***) was a battery of seven subtasks covering the aspects of basic literacy as shown in ***Table 1.*** All of the literacy subtasks were administered individually in the target language. Timing for answering questions was applied consistently to ensure the scores were derived under the same times.

[2] See DRIC/UCC (2016), *Complementary Basic Education (CBE) Learners Assessment: Baseline Report for 2015/2016* for a full account of the process of developing the original instruments.

Table 1. Literacy assessment subtasks

Subtask	Skill	Description The student was asked to …	Scoring
Letter sound identification	letter sound knowledge	… pronounce the sounds of 50 letters from the target language in up to one minute. The letters were presented in random order in a grid of five rows of 10 letters each, with a combination of lowercase and uppercase letters. According to the linguistic characteristics of the given language, in some languages, common two-letter digraphs or consonant blends were included.	This subtask was scored for accuracy as the total number of letter sounds identified correctly in one minute over 50 possible sounds. The students were given up to one minute to sound out as many letters as they could, but excess time was not recorded for students who finished in under one minute. Therefore, only accuracy (as opposed to rate) was scored and reported.
Phonemic awareness	phonemic awareness of word – initial sounds	…provide a word that began with each of 10 phonemes from the target language. The phonemes were presented orally one at a time by the enumerator. The subtask began with an example so that the student could understand what to do. This subtask was allotted five minutes.	This subtask was scored as the number of items correct over 10 possible. Any word beginning with the given sound was counted correct.
Familiar word identification	word decoding and recognition in isolation	… read aloud a list of 20 familiar words (i.e. words likely to be encountered frequently in a school setting) in two minutes. The words were presented in random order on a grid of four rows of five words each. The words consisted of one to three syllables, with two-syllable words being the most common.	This subtask was scored for accuracy as the total number of words read correctly in two minutes over 20 possible. As with letter sound identification, rate adjustments were not calculated and only accuracy was scored and reported.
Reading comprehension (*from silent reading*)	comprehension	… read silently a short, grade-level passage in the target language, then respond orally to four questions (two direct and two inferential) that the enumerator asked based on the passage. The passage was 50 words long. Three minutes were allotted to the silent reading, and five minutes to the comprehension questions.	This subtask was scored as the number of questions answered correctly over four possible.
Oral passage reading fluency	word decoding and recognition in connected text	… read aloud a short, grade-level passage in the target language in one minute.	This subtask was scored for accuracy as the total number of words read correctly in one minute. As with letter sound identification, rate adjustments were not calculated and only accuracy was scored and reported.
Word writing	letter sound knowledge, letter formation, and word encoding	… write five words in the target language one at a time as dictated orally by the enumerator. The words selected for this subtask were familiar words of one to three syllables in length. The time allotted to this subtask was two minutes.	This subtask was scored as the total number of words written correctly over five possible. Words were counted as correct if they were spelled either correctly or phonetically.
Creative writing/ sentence formation	word encoding, language knowledge	… write two meaningful sentences in the target language about the food s/he likes best. The time allotted to	This subtask was scored over two marks possible.

Subtask	Skill	Description The student was asked to …	Scoring
	(vocabulary and grammar)	this subtask was five minutes.	

1.3 Numeracy assessment instrument

Like the literacy assessment, the numeracy assessment (see sample in ***Annex 3***) was administered in the target language. This instrument had a total of nine subtasks. The first three subtasks (number identification and missing number) were administered individually, and the remaining subtasks (problem-solving) were administered as a group to the whole class at the same time; the mechanical problems were presented in writing on paper, and the word problems were read aloud by the enumerator. The students marked their answers on paper. The numeracy subtasks used are described in ***Table 2.***

Table 2. Numeracy assessment subtasks

Subtask	Skill	Description The student was asked to …	Scoring
Number identification: One-digit	knowledge of single-digit number names	… orally identify 50 single-digit numbers from a written list in one minute.	This subtask was scored for accuracy as the total number of numbers correctly identified over 50 possible.
Number identification: Two-digit	knowledge of two-digit number names	… orally identify 40 two-digit numbers from a written list in one minute.	This subtask was scored for accuracy as the total number of two-digit numbers correctly identified over 40 possible.
Missing number	the ability to discern and complete number patterns	… orally identify the missing number to complete a given sequence. The student was presented with five incomplete number sequences on paper. The time allotted to this subtask was three minutes.	This subtask was scored for accuracy as the total number of missing numbers correctly identified out of five possible.
Problem solving: One-digit addition	knowledge and application of basic addition facts; the ability to interpret a situation, make a plan, and solve the problem	… solve and answer in writing two mechanical problems requiring one-digit addition in approximately three minutes and one word problem with one-digit addition in approximately three minutes.	Each subtask was scored for accuracy as the total number correct out of the total number possible.
Problem solving: One-digit subtraction	knowledge and application of basic subtraction facts; the ability to interpret a situation, make a plan, and solve the problem	… solve and answer in writing two mechanical problems requiring one-digit subtraction in approximately three minutes and one word problem requiring one-digit subtraction in approximately three minutes	
Problem solving: Two-digit addition	knowledge and application of basic addition facts to more complex addition problems	… solve and answer in writing four mechanical problems requiring two-digit addition in approximately seven minutes.	

Subtask	Skill	Description The student was asked to …	Scoring
Problem solving: Two-digit subtraction	knowledge and application of basic addition facts to more complex subtraction problems	… solve and answer in writing four mechanical problems requiring two-digit subtraction in approximately seven minutes.	
Problem solving: Multiplication	knowledge and application of multiplication facts	… solve and answer in writing two mechanical problems requiring multiplication in approximately four minutes and one word problem with multiplication in approximately three minutes.	
Problem solving: Division	knowledge and application of division facts	… solve and answer in writing two mechanical problems requiring division in approximately three minutes and one word problem requiring division in approximately three minutes.	

1.4 Recruitment, selection, and training of enumerators

For this endline study, the study's two local evaluation partners, JEAVCO Associates and PAB Development Consultants, relied on CBE baseline study data collectors. This was done in order to minimise the amount of additional training needed and to use data collectors familiar with the CBE study for improved data quality. Mr Emmanuel Baapeng (responsible for initial recruitment of baseline data collectors) was tasked with ensuring that all baseline data collectors were once again available for this round of data collection. In all, only one person was not recalled for the endline study – due to unsatisfactory performance during the baseline study. A new, substitute data collector was trained alongside the remaining data collectors. As with the baseline, the training was undertaken at the M&J Hospitality in Tamale. (See ***Annex 4*** for a report on the training of data collectors.)

1.5 Management of data collection

The endline data collection management procedures were the same as those used for the baseline study:

1. The data collection teams were led by staff from JEAVCO Associates and PAB Consulting – a National Quality Assurance Lead and a Data Collection Specialist.
2. The data collectors were subdivided into 14 teams. The size of the teams depended on the number of pupils in a specific language group and the spread of the participating districts. For example, one Dagbani language group had a team of five enumerators, while the Likpakpaln language group had a team of six enumerators; these were the largest group sizes. In the case of the Gurune and Kasem language groups, each team consisted of a single data collector in a single district. In Ashanti Region, there was only one participating district and so one enumerator was assigned to the district (see ***Annex 5***). Each language team had a leader who remained in frequent contact with the entire data collection team and also the team leaders from PAB and JEAVCO.
3. Prior to the deployment of teams to the field, the IPs were alerted by letters from the Project Management Unit. Where possible, actual dates and times of visits by the data collection teams were communicated to the IPs via telephone calls from team leaders.

4. Throughout the data collection process, work was overseen by the National Quality Assurance Lead and the Data Collection Specialist. Regular communication was maintained throughout the process using a 'WhatsApp' group for all data collectors and team leaders.
5. A tablet-based app which was developed for entering data collected during the baseline study was updated with the endline instruments. As the teams visited the schools in the local areas and enumerators assessed children within their schools, data were captured on the tablets and relayed electronically to a service centre. The data ultimately were edited and then analysed for report writing.

1.6 Component scores and proficiency levels

As previously noted, the literacy assessment used for this study consisted of seven subtasks, while the numeracy assessment was made up of nine subtasks. While it is possible to examine each subtask individually, it is often more valuable to combine subtasks to calculate competencies or student profiles or proficiencies. In the simplest sense, total scores can be combined across subtasks in order to create an unweighted composite score. This has the benefit of simplicity but does not take into account the relationships among subtasks or their relative importance in terms of the overall test scale. Two of the most common approaches for creating composite scaled scores are item response theory (IRT) and principal component analysis (PCA).

At baseline, composite scores were initially created using both approaches and similar results were obtained. Since item response theory relies on more complicated modelling decisions (and greatly benefits from the availability of item-level responses – which were not entered for these data), the more simplified PCA approach was selected for producing all composite scores. The basic premise behind this approach is that a large number of correlated observed variables (e.g. reading subtasks) can be reduced to a smaller set of independent composite variables (principal components). In other words, instead of examining each subtask individually, it is possible to use the relationships among subtasks to create composite scores that were based on the underlying constructs they measured. This is not to say that individual subtask scores were not examined at all but simply that the main focus of this report (as with the baseline report) was on understanding reading and mathematics competencies, which are most easily examined using composite scores.

The same composite scores that were used in the baseline were also used in this endline evaluation (for direct comparability). For both the literacy and mathematics assessments, composite scores were created for basic and advanced skills. In the end, five separate score categories were defined ***(Table 3).*** While exploratory analyses showed that it would be possible to retain just one component for overall scores for both maths and reading, splitting the tests into sub-components showed only small reductions in test scale reliability (due to the small reduction in items) but increased the amount of variation explained by the first component for each group, which was used to create component scores. All Cronbach's alphas in Table 3 are well above the acceptable cut-off of 0.70. Additionally, the final column of the table shows that between 70% and 91% of the variation in scores was explained by the categories as defined in these models. Therefore, the number of subtasks was effectively reduced for analyses, while still achieving variation (as opposed to just a single measure).

Table 3. Component score categories, by subtask

Component score category	Subtasks	Cronbach's alpha (internal consistency)	Proportion of variance explained by first component
Basic Reading	• Letter sound identification • Phonemic awareness	0.90	0.91
Advanced Reading	• Familiar word identification • Reading comprehension	0.89	0.90
Writing	• Word writing	0.88	0.89

Component score category	Subtasks	Cronbach's alpha (internal consistency)	Proportion of variance explained by first component
	• Creative writing/sentence formation		
Basic Numeracy	• Number identification • Missing number • One-digit addition • One-digit subtraction	0.89	0.70
Advanced Numeracy	• Two-digit addition • Two-digit subtraction • Multiplication • Division	0.90	0.78

After the component scores were created, all scores were scaled from 0 to 100, for ease of interpretation.

As a final step, the scaled component scores were then divided into four proficiency or competency levels. These were defined based on the students' performance on the component scores, as follows:

1. **Non-performer,** comprising those who scored zero on a component score;
2. **Beginner,** comprising those who scored greater than zero but less than 50;
3. **Approaching proficiency,** comprising those who scored greater than 50 but less than 80; and
4. **Proficient,** comprising those who scored greater than 80.

While it ultimately would have been possible to create any number of proficiency categories, experience from prior work has shown that using four or five categories produces easily interpretable results while still providing enough variation to measure growth across categories. Zero scores represent a natural category (and are an important focus), as do scores reaching high standards (typically mastery of a subtask would be related to scores of 80–90%). Since this CBE research focuses to some degree on reducing zero scores, it also was important to have a category to represent movement out of this group. However, there is a clear difference between those just above zero and those just below the highest threshold – hence the final decision to divide the middle group into two categories.

Since these are scaled composite scores, the numbers themselves do not have easily interpretable independent meaning. They do not represent percent correct, but instead are simply a factor-based composite that has been scaled from 0 to 100 (as the actual scale of approximately −1 to 3.5 would be too obscure for most audiences). Therefore, the task was to create reasonable categories that could be understood by a variety of stakeholders. It is important to keep in mind that 'proficiency' is a relative term and that these cut-points are not intended to be used as standards for the CBE programme. Instead, they are designed to create a baseline against which growth in student performance can be measured.

That being said, these categories can be examined in relation to subtask performance in order to determine their appropriateness. The average percent correct for each subtask (for each of the basic and advanced reading category score categories) is displayed in ***Table 4***. Note that these averages remained consistent from baseline to endline, showing that these categories continued to measure the same constructs and were therefore comparable over time.

Table 4. Average percent correct for reading subtasks

Subtask	Basic Reading Score (in %)				Advanced Reading Score (in %)			
	Non-Performer	Beginner	Approaching Proficiency	Proficient	Non-Performer	Beginner	Approaching Proficiency	Proficient
	0	>0; <50	≥50; <80	≥80	0	>0; <50	≥50; <80	≥80
Letters – Baseline	0	21.6	62.2	88.8				
Letters – Endline	0	23.8	66.4	93.7				
Phonemic awareness – Baseline	0	19.0	64.1	90.1				
Phonemic awareness – Endline	0	23.7	63.6	93.3				
Familiar words – Baseline					0	24.0	65.2	87.6
Familiar words – Endline					0	27.6	69.0	92.5
Reading comprehension – Baseline					0	23.0	61.0	91.8
Reading comprehension – Endline					0	21.1	59.6	93.4

As expected, Table 4 shows that students scoring zero (i.e. non-performers) in both the basic and advanced reading categories were unable to answer a single item correctly on the respective subtasks that make up the composite scores. Those in the beginner categories (score up to 50) were able to answer about one-fifth to one-quarter of the questions correctly (showing they were just starting out in the reading process). Those who were approaching proficiency (50–80) averaged about 60–70% correct (meaning that based on most standards for mastery, they were not scoring high enough to be considered proficient in those skills). For those who were proficient (≥80), the averages were approximately 90% (which is a high but reasonable standard for claiming that students have mastered the skills and are ready to move on to the next, higher-order skills). While arguments could be made about proficiency/mastery not needing to be as high for certain subtasks, it was decided that it would be highly preferred to have consistent cut-points across scales for ease of interpretation and reading (once again keeping in mind that these scores are meant to be estimates, as opposed to targets or standards).

Lastly, the four categories offer intuitive and easily interpretable results for all audiences: Non-performers (unable to answer a single question correctly); beginners (with a score above zero but less than 50 – i.e. able to answer some questions correctly but unable to obtain a score of more than half the scale); approaching proficiency (with a score of at least 50 but still below proficiency); and proficient (with a score of 80 or above – which is a high but attainable upper threshold). Composites were created using an approach similar to how it is done for any other index (much like the Program for International Student Assessment [PISA], Trends in International Mathematics and Science Study

[TIMSS], etc.). Accordingly, there is nothing in particular that says that proficiency needs to be 80[3]. These cut-points were mainly chosen for ease of interpretation – 0, 50, and 80 are relatively intuitive numbers in terms of 'can't read', 'can read some', and 'seems to be doing really well'.

2. The CBE Cycle 4 Endline Sample

2.1 Sample of students

The original intention was to follow the 2,401 students who were assessed at baseline and to administer the assessments to them again at endline. These students were initially selected for the baseline learner assessment sample using a stratified random sampling approach intended to provide proportional representation of the sample by gender, language, region, district and centre. Despite the best efforts by the data collection team, however, it was ultimately not possible to test all 2,401 students at endline, due mostly to dropout from the CBE programme. In the end, 399 students from the baseline sample were left out of the endline sample, bringing down the total sample to 2,002 students (17% sample attrition). Note that attrition throughout this section refers simply to students who were unavailable at endline but no claims are made about the reason for this, as it could be accessibility, absenteeism, or dropout. In the next phase of the tracker study for Cycle 4, information will be gathered about actual dropout rates, at which time additional analyses will be conducted.

Since it was not possible to follow 399 students from the baseline sample, it is essential to determine how these cases should be dealt with. On the one hand, if the students who were unavailable at endline had baseline results similar to their counterparts who were tracked, and if there were no systematic differences between dropout students (particularly by IP or CBE centre), an argument could be made for averaging results across the full samples. If, on the other hand, there was differential attrition (i.e. systematic differences in those who dropped out, particularly by IP or CBE centre), the analyses must account for it. Additionally, since some of the analyses in this report were designed to examine improvement at the student level, it was necessary to limit the sample to those students who were available at both time points.

While the overall attrition rate for the sample was 17%, the attrition rates for each IP are shown in ***Table 5.*** This table shows that there was variation in attrition by IP, with School for Life retaining more than 90% of their baseline sample, while Action Aid and GILLBT both had attrition rates of greater than 25%.

Table 5. Sample attrition, by implementing partner

Partner	Percentage of students available only at baseline	Percentage of students available at baseline and endline
Action Aid	27.3%	72.7%
AfriKids	12.5%	87.5%
CARE International	24.1%	75.9%
Ghana Institute of Linguistics, Literacy and Bible (GILLBT)	28.3%	71.7%
IBIS Ghana	15.6%	84.4%
Link Community	18.0%	82.0%

[3] It is important to note that the number 80 on the scale is not the same as 80% correct. Putting it on a scale of 1-100 makes it somewhat easier to interpret but should not be interpreted as a percentage score.

Partner	Percentage of students available only at baseline	Percentage of students available at baseline and endline
Development (LCD)		
Plan Ghana	21.5%	78.5%
ProNet	11.1%	88.9%
School for Life	9.5%	90.6%
World Education	18.2%	81.8%
Total	**16.6%**	**83.4%**

The variation in attrition rates in IPs is modest, but it hides some of the greater variation in centres – which ranged from 0% to 100%. The one centre with an attrition rate of 100% was Wulugu (which was inaccessible during the endline data collection due to severe flooding). Therefore, Wulugu was removed from all analyses, since no students were assessed at endline. There were, however, 11 centres with exceedingly large attrition rates (between 40% and 50%) and 18 centres without any attrition at all.

Annex 6 gives details of the sample sizes across the IPs, regions, districts and languages as well as the number and names of the centres per district. The CBE Cycle 4 baseline report provides a complete overview of the baseline sample by characteristics from the child questionnaire.

2.2 Attrition analysis

In order to examine differences between those students who dropped out of the sample and those who remained through the endline assessment, ***Table 6*** provides the results of two-sample *t*-tests that were run on baseline scores for those students who were and were not able to be tested at both time points.

Table 6. Baseline mean percent scores by follow-up status

Subtasks	Available only at Baseline; Baseline % Scores	Available at Baseline and Endline; Baseline % Scores
Literacy Assessment Subtasks		
Letter sound identification	22.6	30.9**
Phonemic awareness	23.6	29.2**
Familiar word identification	18.1	20.7
Reading comprehension	15.2	20.6**
Word writing	15.8	19.8**
Creative writing/sentence formation	9.7	14.8**
Numeracy Assessment Subtasks		
Number identification: One-digit	41.3	49.6**
Number identification: Two-digit	28.0	33.1**

Subtasks	Available only at Baseline; Baseline % Scores	Available at Baseline and Endline; Baseline % Scores
Missing number	34.6	42.3**
One-digit addition	49.3	57.4**
One-digit subtraction	43.0	48.7**
Two-digit addition	22.1	29.9**
Two-digit subtraction	16.4	24.3**
Multiplication	28.4	31.4
Division	27.3	31.2*

* $p < 0.10$; ** $p < 0.05$.

These results show that students who dropped out (or who were unavailable for follow-up) were statistically significantly lower performers than those who were available at both time points. This result was found to be true for all but two subtasks (familiar word identification and multiplication). It is not surprising that students who dropped out of the programme tended to be those who were underperforming at the start of the school year.

In addition to analysis of literacy and numeracy performance, the proportion of students who dropped out of the sample was examined by key characteristics of interest from the baseline sample (shown in ***Table 7***)**.** There was no significant difference in dropout rates by sex, work outside the home or home activities that involved reading/writing. Prior school attendance was marginally significant, with previous school attendance slightly increasing the likelihood of attrition. The largest difference came in terms of absenteeism. When asked at baseline whether or not they had been absent in the prior five days, those who said 'yes' were significantly more likely to have dropped out of the sample than those who were not absent (18% v. 14%). This suggests that attrition rates were random for many characteristics but that those who had previously been to school and/or showed greater levels of attendance at baseline were also more likely to be available for follow-up at endline.

Table 7. Educational background, by attrition

Questionnaire item	Dropout rate (%)
What is the pupil's sex?	
Female	16.6
Male	15.0
Have you ever been to school?	
No	14.7*
Yes	19.4*
Do you do any work outside of home?	
No	16.6
Yes	14.8
Have you been absent in the last five days?	
No	13.7**
Yes	17.8**

Questionnaire item	Dropout rate (%)
Do you do any activity at home that involves reading or writing?	
No	16.6
Yes	15.4

* $p < 0.10$; ** $p < 0.05$.

3. Results

This chapter presents the results of the endline assessment. Throughout the initial analyses, the decision was made to provide the endline estimates as well as the estimates of learning gains. This approach gives both a measure of the state of literacy and numeracy in the CBE centres, and an estimate of how much students had improved since the start of the school year.

Table 8 shows the overall mean percent scores on the baseline and endline literacy and numeracy assessments,[4] as well as the change in scores between the two time points. Overall, there was a significant increase in the mean scores for students across all 15 of the subtasks. Students continued to underperform on the literacy subtasks (as compared with numeracy) but the average gains in literacy were higher. On average, mean literacy scores increased at a nearly consistent 25 percentage points (ranging from 24 percentage points in creative writing and reading comprehension to 28 percentage points in familiar word identification). The average increase in numeracy scores was somewhat more varied, with the smallest gains coming in missing number and one-digit addition, and the largest gains coming in multiplication and division.

Table 8. Overview of mean percent scores (baseline, endline and growth)

Subtasks	Baseline mean percent score (%)	Endline mean percent score (%)	Change in mean percent score (percentage points)
Literacy Assessment Subtasks			
Letter sound identification	30.9	57.1	26.1
Phonemic awareness	29.2	56.2	26.9
Familiar word identification	20.9	48.9	28.0
Reading comprehension	20.7	44.8	23.8
Word writing	19.9	45.2	25.5
Creative writing/sentence formation	14.9	38.6	23.7
Numeracy Assessment Subtasks			
Number identification: One-digit	49.8	70.1	20.2
Number identification: Two-digit	22.3	54.7	21.5
Missing number	42.4	59.8	17.2
One-digit addition	57.8	77.5	19.5
One-digit subtraction	48.7	68.9	20.1

[4] Note that baseline scores in this report may differ slightly from scores in the CBE Cycle 4 baseline report due to the reduced sample based on attrition.

Subtasks	Baseline mean percent score (%)	Endline mean percent score (%)	Change in mean percent score (percentage points)
Two-digit addition	29.9	57.4	27.4
Two-digit subtraction	24.3	52.0	27.6
Multiplication	31.4	59.5	28.0
Division	31.2	60.0	28.9

In addition to mean percent scores, evaluators are often interested in zero scores (i.e. the percentage of students who are unable to correctly answer a single item on a given subtask). Examining both increases in mean percent scores and decreases in zero scores allows for a more complete understanding of average learning gains as well as improvements for students at the lowest end of the achievement spectrum. Since one of the main focuses of the CBE programme is to provide out-of-school youth with a preparatory experience for traditional primary schooling, it is essential that CBE centres be able to support the lowest-performing students. Zero score estimates from the baseline and endline assessments, as well as a measure of 'reduction' in zero scores, are shown in ***Table 9***.

Table 9. Overview of mean percent zero scores (baseline, endline and reduction)

Subtasks	Baseline percent zero scores (%)	Endline percent zero scores (%)	Change in percent zero scores (percentage points)
Literacy Assessment Subtasks			
Letter sound identification	15.1	5.3	–9.8
Phonemic awareness	34.6	14.5	–19.9
Familiar word identification	59.2	21.9	–37.1
Reading comprehension	65.0	33.3	–31.2
Word writing	62.3	30.2	–32.0
Creative writing/sentence formation	78.1	47.9	–30.1
Numeracy Assessment Subtasks			
Number identification: One-digit	8.0	2.7	–5.4
Number identification: Two-digit	20.8	9.9	–10.9
Missing number	21.2	9.5	–11.5
One-digit addition	21.2	7.8	–13.2
One-digit subtraction	29.3	12.2	–17.2
Two-digit addition	46.0	16.9	–29.0
Two-digit subtraction	54.3	23.0	–31.1
Multiplication	47.7	20.9	–26.8
Division	51.4	21.2	-30.2

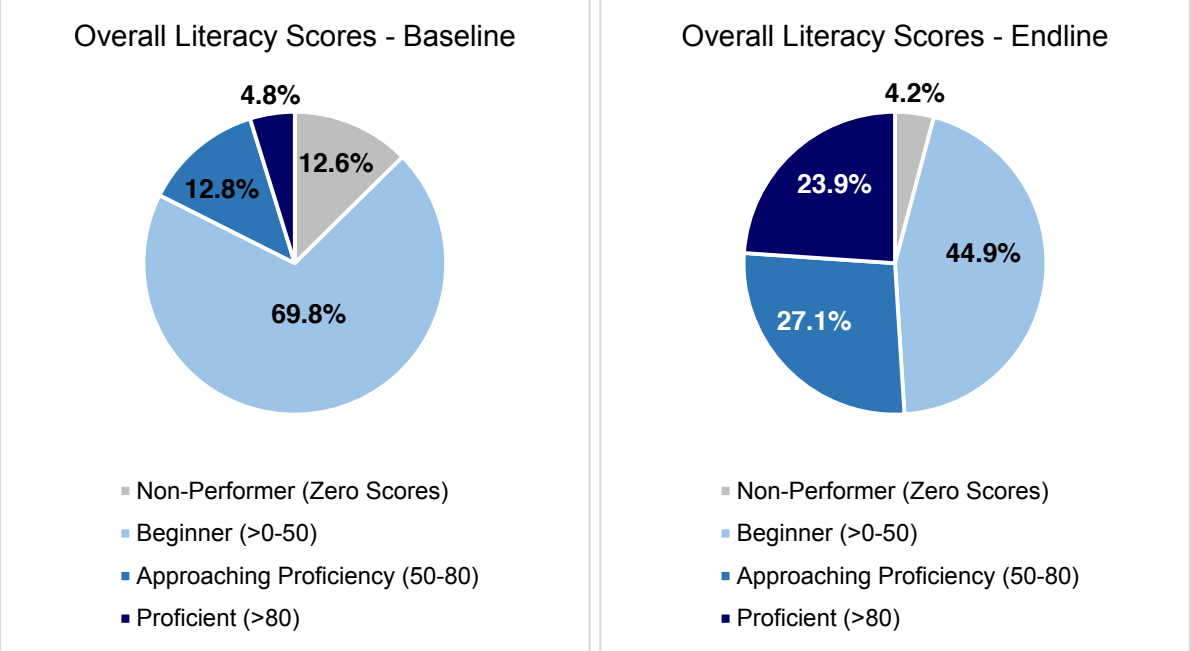
While the expectation would be for any learning programme to reduce zero scores over the course of a nine-month period, Table 9 shows impressive reductions across all subtasks. At endline, the percentage of zero scores was significantly reduced for all literacy subtasks. The largest absolute reduction was in familiar word identification (showing that students significantly increased their word recognition and decoding skills), although the largest relative reduction was for letter sound identification (bringing the total percentage of zero scores at endline down to only 5%). For mathematics, fewer than 10% of students were unable to answer a single basic question (one- and two-digit number identification, missing number and one-digit addition) and no more than a quarter of students were unable to correctly answer a single math question on any subtask (with the highest zero score proportion coming in two-digit subtraction – though that skill did also have the largest overall reduction in zero scores).

In addition to mean percent gains or reductions, individual student gains were also calculated (but not shown in the tables above). Specifically, a student was categorised as having improved if s/he was able to answer at least one more question correctly on the endline as compared to the baseline. This is notably an over-simplified metric but it is interesting to note that while more than half of the students were able to increase their literacy scores from baseline to endline on the first five subtasks (with 78% increasing their scores on letter sound identification), only 40% were able to do so on creative writing. With regard to numeracy, at least half of the students improved in all skills except for single-digit addition (46%) and single-digit subtraction (49%), with the largest proportion of increases coming from single-digit number identification (70%).

3.1 Overall proficiency levels

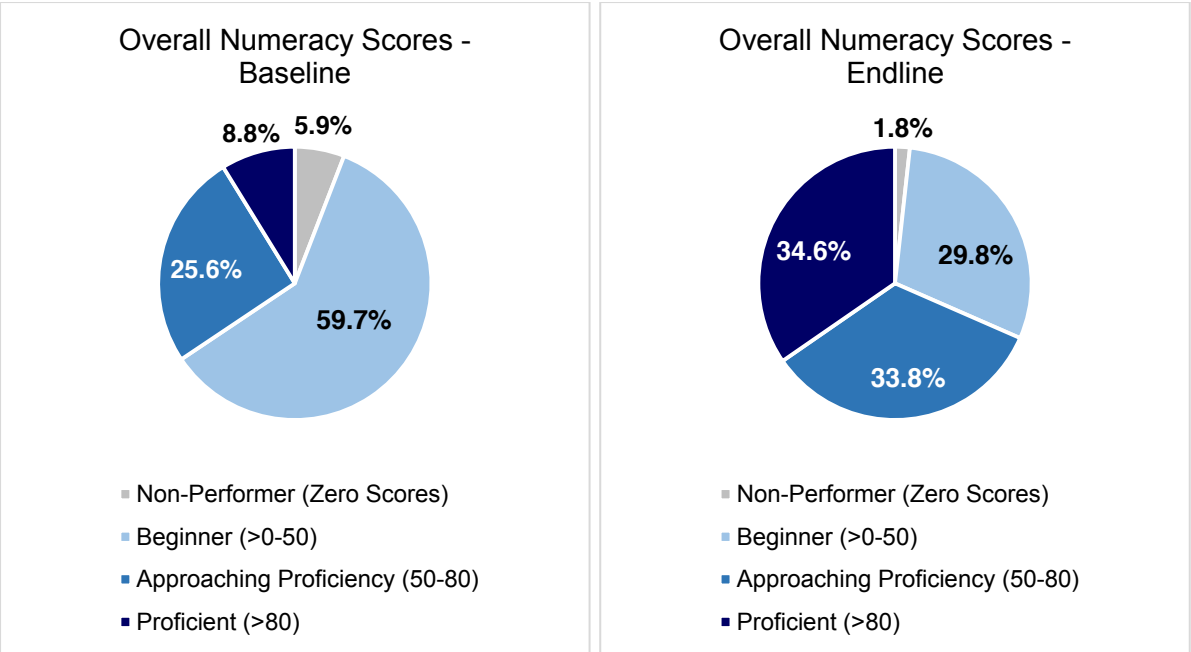
The overall proportions of students falling into each proficiency level for literacy at baseline and endline are shown in ***Figure 1.*** The figure illustrates large improvements across all categories during the nine-month programme. This is seen by the reductions in the lowest two categories and the increases in the upper two categories. More specifically, while 12.6% of students were non-performers in literacy at baseline, that number was cut by two-thirds by endline, to 4.2%. The beginning category also saw a large reduction, from approximately 70% at baseline to approximately 45% at endline. The proportion of students in the approaching proficiency category more than doubled, from 12.8% to 27.1%. Perhaps the most impressive finding is that the proportion of students in the proficient category more than quadrupled, from 4.8% at baseline to nearly a quarter (24%) at endline.

Figure 1. Overall literacy scores by proficiency level, at baseline and endline



The overall results for the numeracy categories are presented in ***Figure 2.*** Much like the literacy scores, there were significant improvements in all four categories for the numeracy assessment. There were large reductions in the proportion of non-performers (from 6% to 2%) and the proportion of students in the beginner category was cut in half (from 60% to 30%). While there was only a modest increase in those approaching proficiency (from 26% to 34%), the largest increase was in the proficient category, where the proportion of students rose from 9% at baseline to 35% at endline.

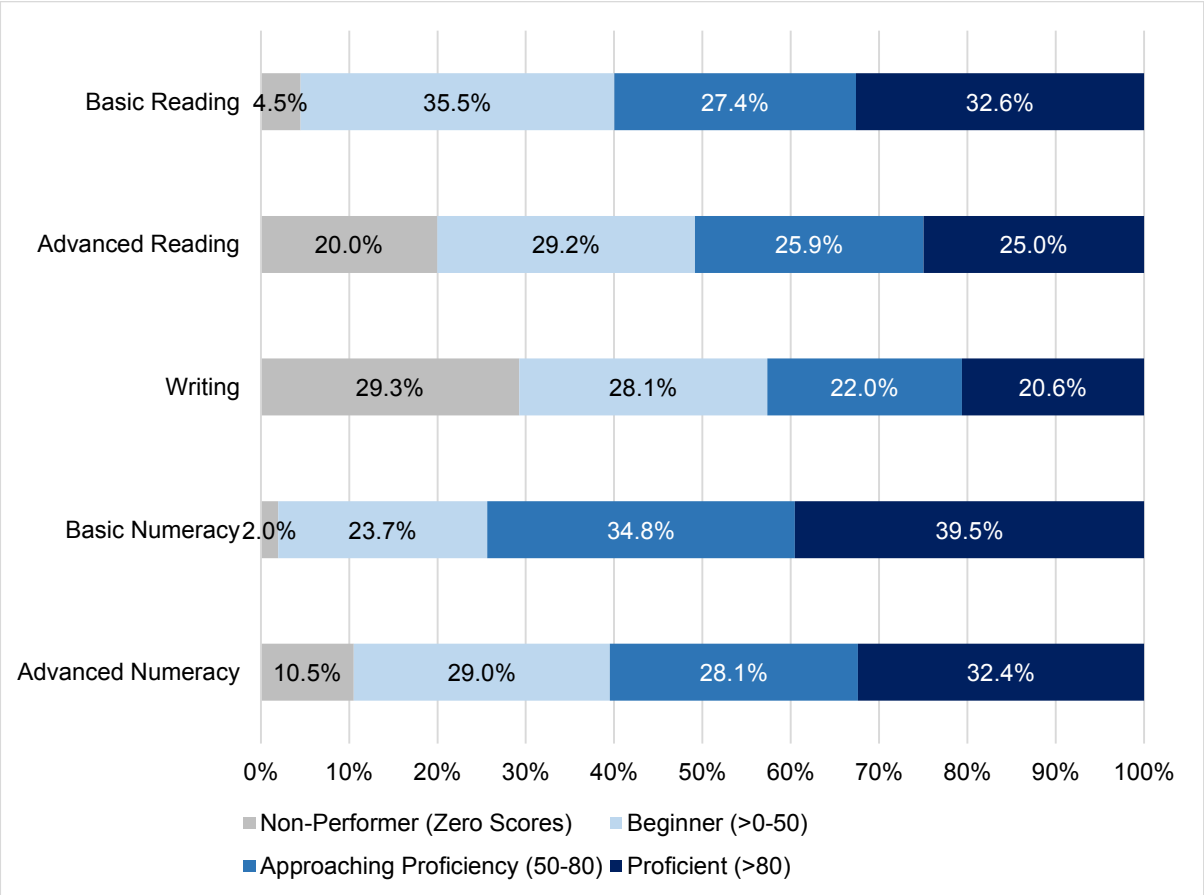
Figure 2. Overall numeracy scores by proficiency level, at baseline and endline



When the literacy scores are broken down into the three component score categories of basic reading, advanced reading, and writing (***Figure 3***), it becomes clear that these students were performing well on basic reading skills (made up of letter sound identification and phonemic awareness), with 60% of students either proficient or approaching proficiency. This represents an increase from only 23% at baseline. In the advanced reading score, 20% of students were non-performers and 29% were beginners, as opposed to baseline, when the proportion of students in the non-performer category was larger than these two profiles combined (57%). Writing continued to prove to be the hardest skill on the assessment, with nearly one-third of students in the lowest category and only one-fifth of students in the highest. However, this was still a marked improvement over baseline, when only 15% of students were in either of the top two categories combined.

Basic numeracy showed the best performance of any category. Only 2% of students were considered to be non-performers, while nearly 40% were found to be proficient in basic numeracy skills (i.e. number identification, missing number, one-digit addition, and one-digit subtraction). The approximately one-quarter of students (26%) in the bottom two categories at endline is an impressive reduction from the 58% that made up those two categories at baseline. Finally, the advanced numeracy profile shows an approximately even distribution of students across three of the profiles, with a smaller percentage in the non-performing category. This, once again, is a large improvement over the baseline, when 33% were non-performers and only 6% were proficient.

Figure 3. Overall component scores by proficiency level, at endline



3.2 Differences by gender

Male students slightly outperformed female students on the endline assessment. These results are displayed in ***Figure 4***. It is clear from this figure that baseline scores for males and females were

nearly identical for both overall literacy and overall numeracy. However, there were slightly larger gains for male students on both assessments—though these gains were not statistically significantly different.

Figure 4: Overall performance for males and females, at baseline and endline

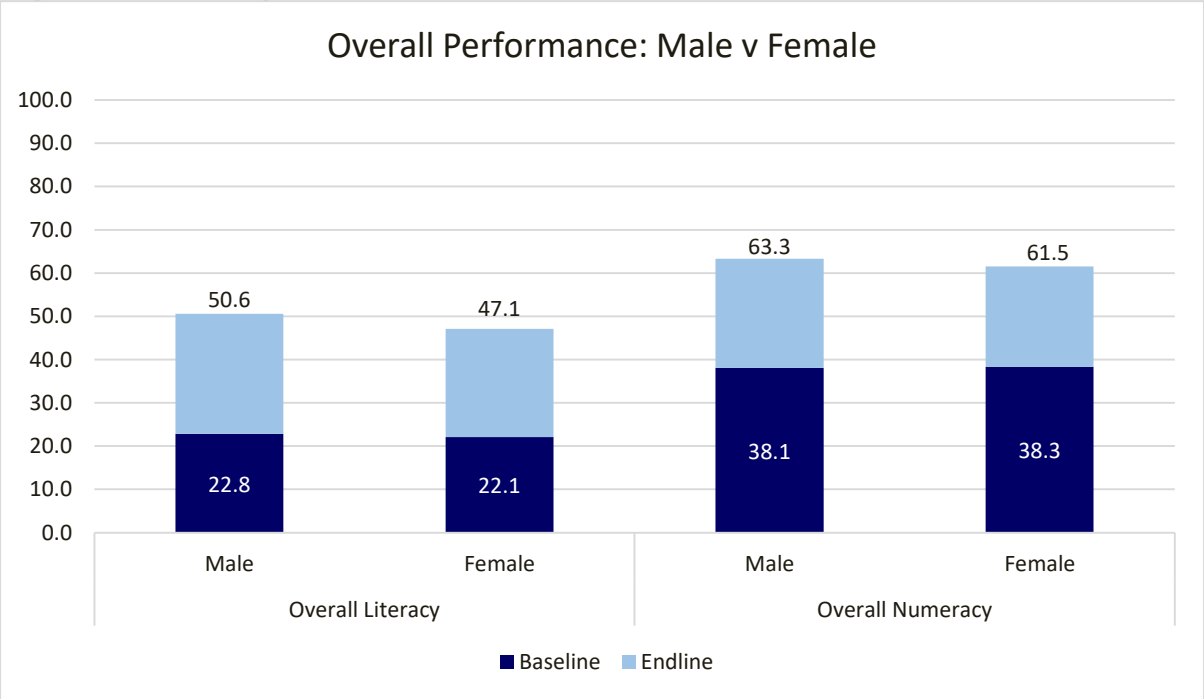
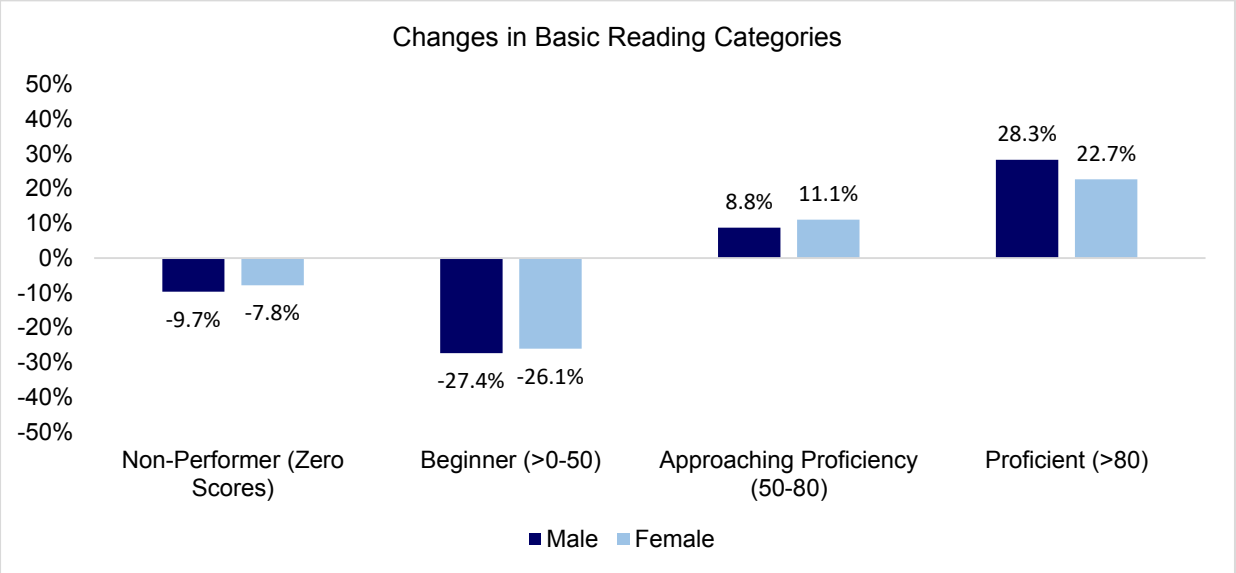
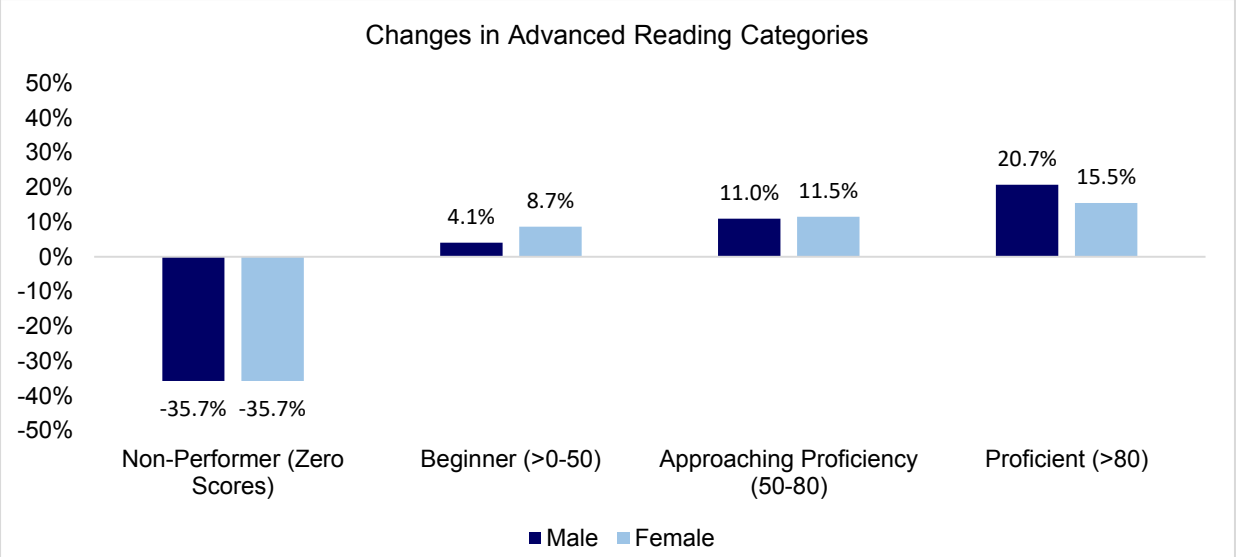
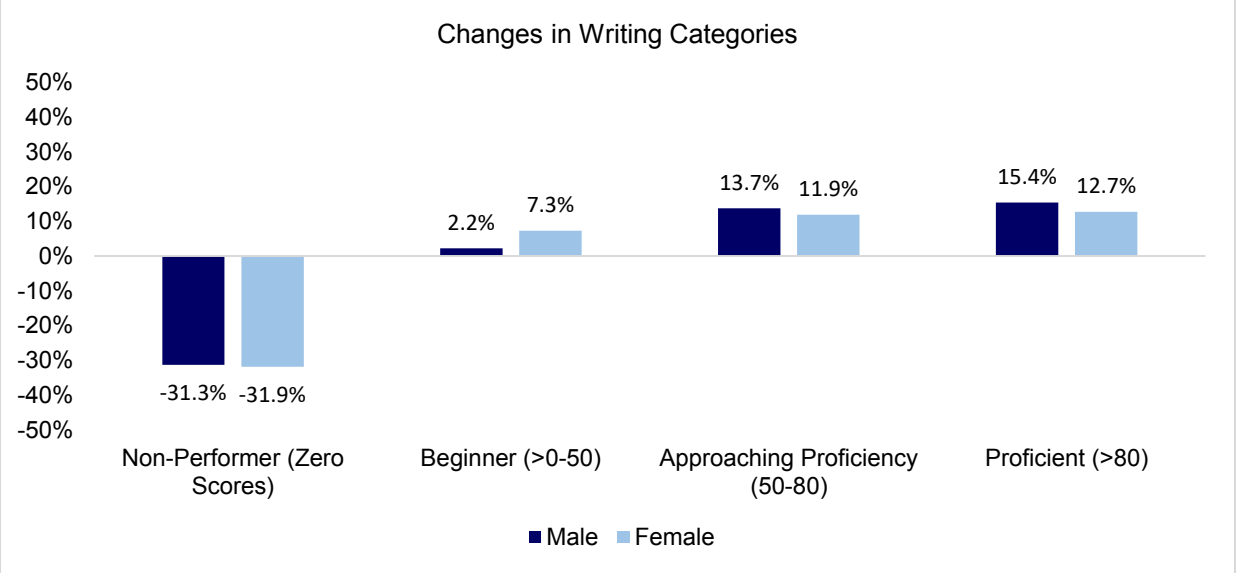


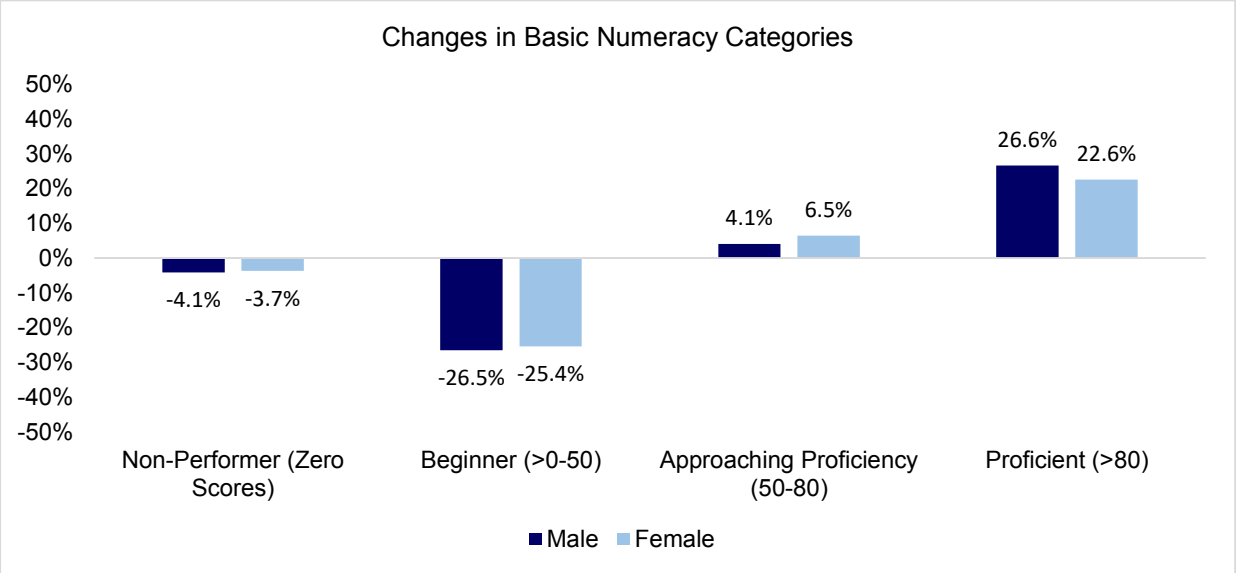
Figure 5 shows that the higher endline performance was due at least in part to a greater proportion of male students moving into the proficient category. This finding held for all five of the component scores. Figure 5 shows similar reductions in non-performers for boys and girls in all of the competencies (reading, writing, and numeracy). However, the overall shift for girls was more modest, with a smaller proportion of girls moving up to proficiency from the lower categories.

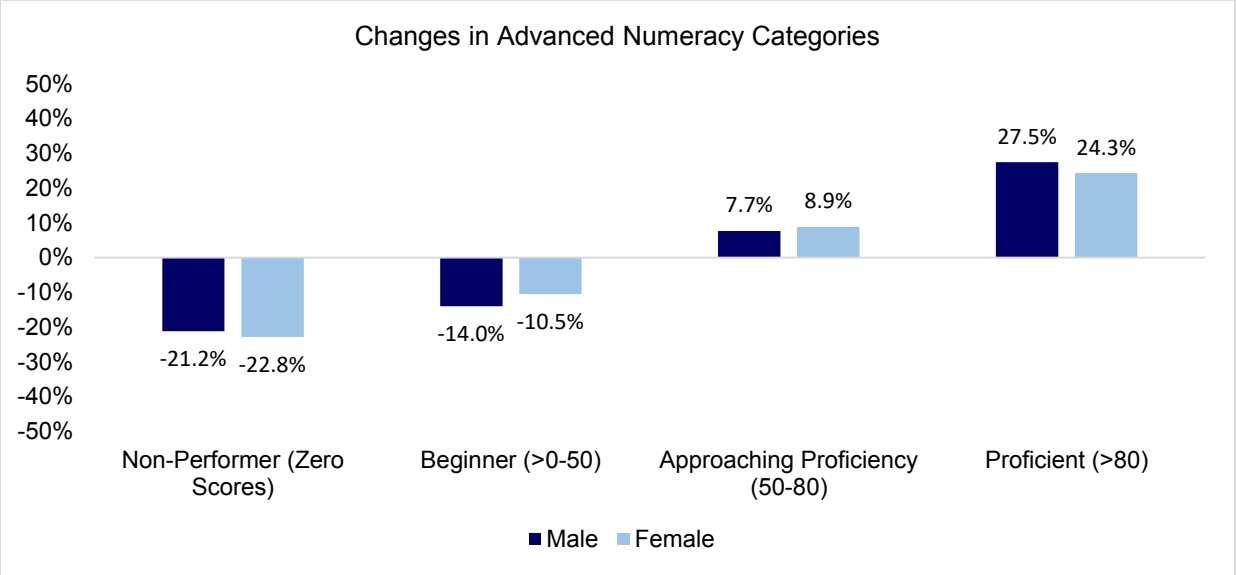
Figure 5. Component score changes, by proficiency level and gender




Changes in Advanced Reading Categories
50%
40%
30%
20%
10%
0%
-10%
-20%
-30%
-40%
-50%
-35.7%
-35.7%
4.1%
8.7%
11.0%
11.5%
20.7%
15.5%
Non-Performer (Zero Scores)
Beginner (>0-50)
Approaching Proficiency (50-80)
Proficient (>80)
Male
Female


Changes in Writing Categories
50%
40%
30%
20%
10%
0%
-10%
-20%
-30%
-40%
-50%
-31.3%
-31.9%
2.2%
7.3%
13.7%
11.9%
15.4%
12.7%
Non-Performer (Zero Scores)
Beginner (>0-50)
Approaching Proficiency (50-80)
Proficient (>80)
Male
Female

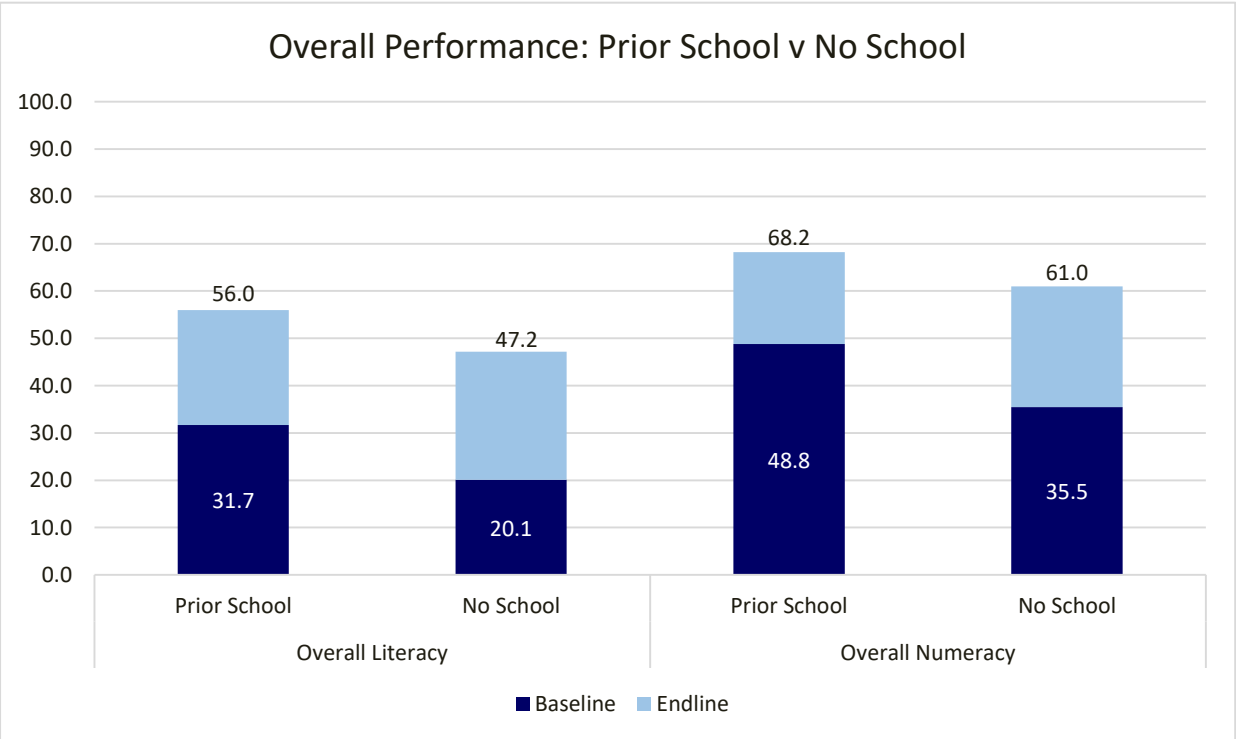

Changes in Basic Numeracy Categories
50%
40%
30%
20%
10%
0%
-10%
-20%
-30%
-40%
-50%
-4.1%
-3.7%
-26.5%
-25.4%
4.1%
6.5%
26.6%
22.6%
Non-Performer (Zero Scores)
Beginner (>0-50)
Approaching Proficiency (50-80)
Proficient (>80)
Male
Female



3.3 Differences by previous school attendance

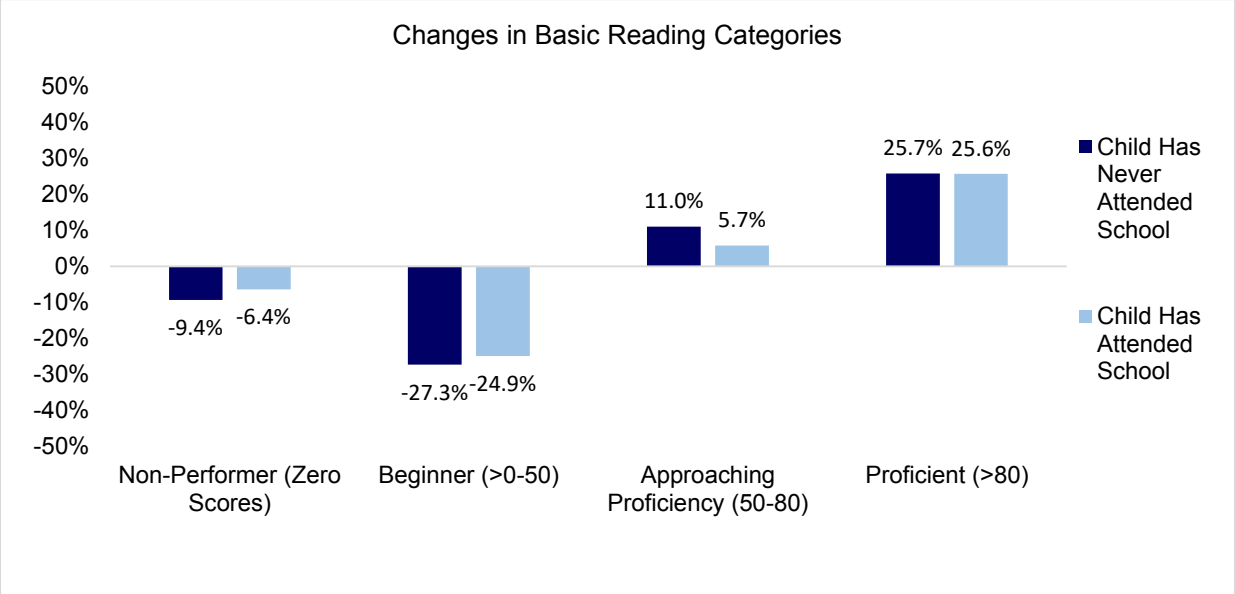
At both baseline and endline, students who had attended school prior to the CBE programme showed higher performance than students who had not. However, those students in the CBE programme without prior school attendance narrowed the gap from baseline to endline, as shown in ***Figure 6***. For example, while students without prior school attendance were 11 points behind their school attending peers in overall numeracy at baseline, the difference was reduced to 7 points at endline.

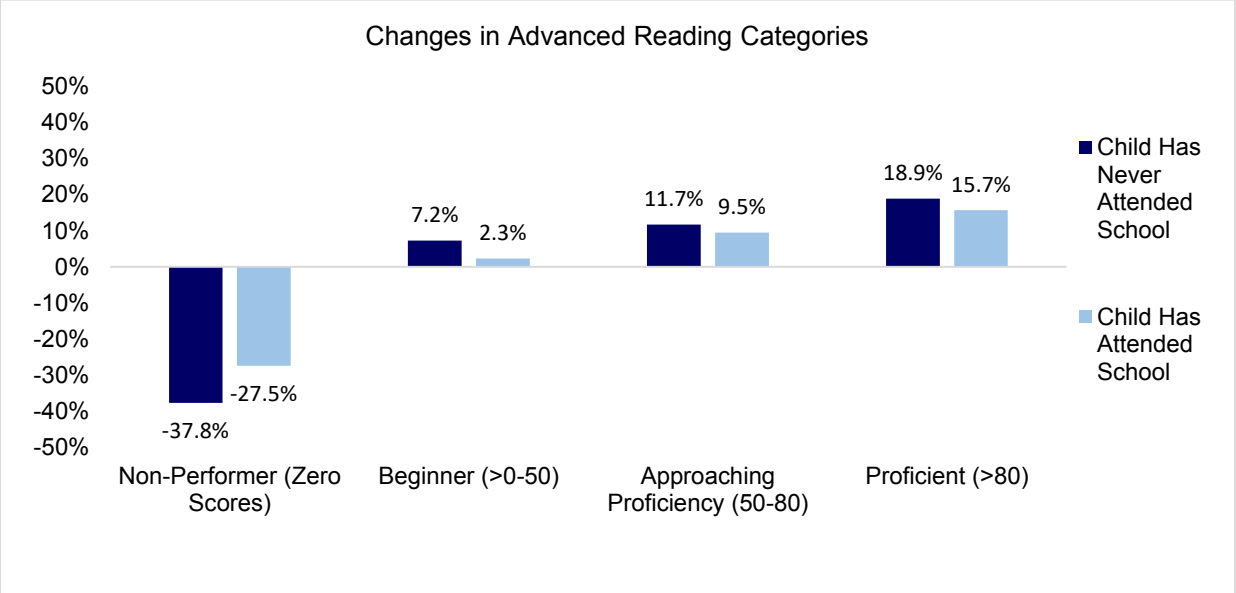
Figure 6: Overall performance by prior school attendance, at baseline and endline

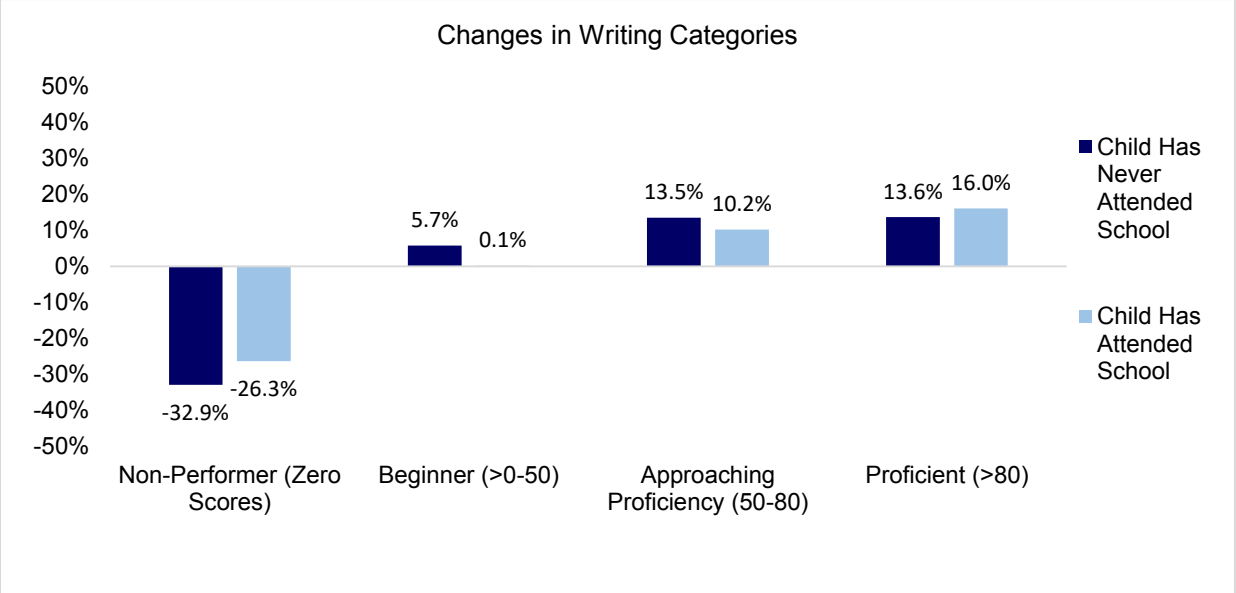


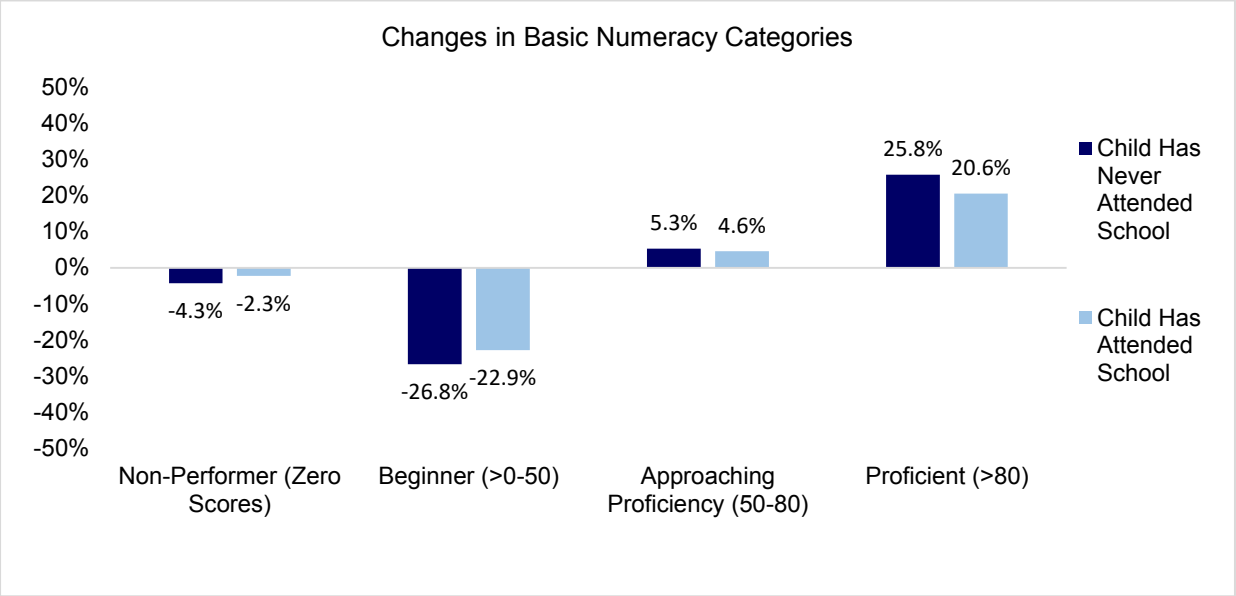
Using the basic reading competency as an example, it is clear that those who never attended school had larger reductions in the proportion of students in the bottom two categories and a larger increase in the proportion approaching proficiency, as shown in ***Figure 7***. The proficient category saw similar increases for both groups. This trend held for each of the five competencies, with a few small exceptions. Overall, the gap clearly was lessening (although an examination of the raw scores showed that a small gap did still exist).

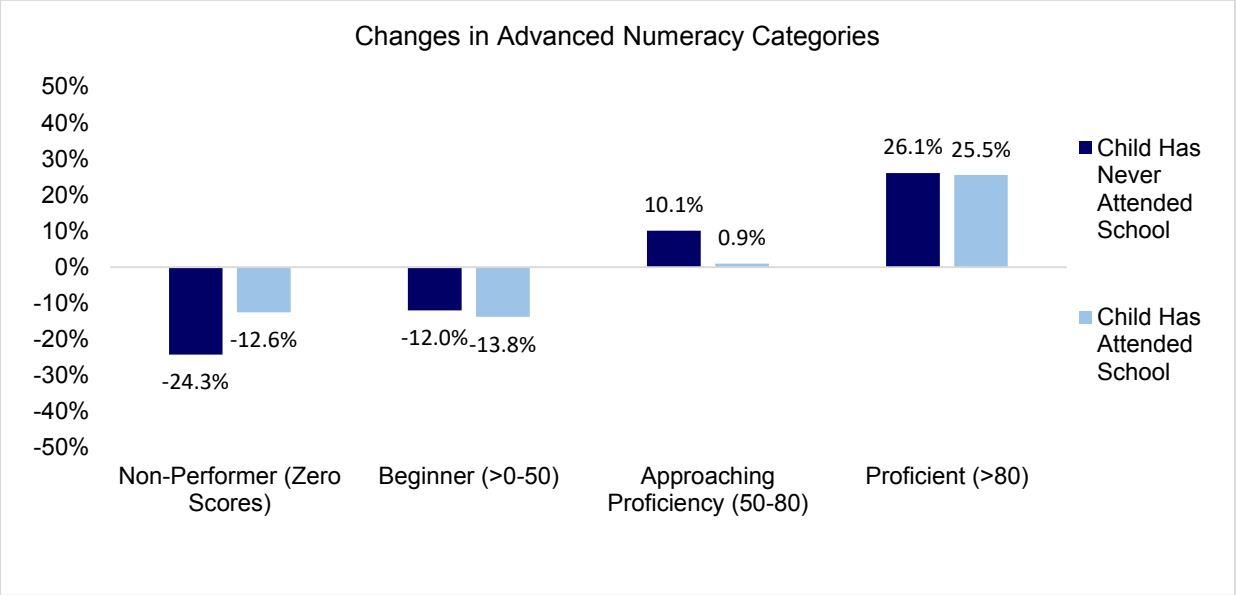
Figure 7. Component scores by proficiency level and previous school attendance






Changes in Writing Categories
50%
40%
30%
20%
10%
0%
-10%
-20%
-30%
-40%
-50%
-32.9%
-26.3%
5.7%
0.1%
13.5%
10.2%
13.6%
16.0%
Non-Performer (Zero Scores)
Beginner (>0-50)
Approaching Proficiency (50-80)
Proficient (>80)
Child Has Never Attended School
Child Has Attended School


Changes in Basic Numeracy Categories
50%
40%
30%
20%
10%
0%
-10%
-20%
-30%
-40%
-50%
-4.3%
-2.3%
-26.8%
-22.9%
5.3%
4.6%
25.8%
20.6%
Non-Performer (Zero Scores)
Beginner (>0-50)
Approaching Proficiency (50-80)
Proficient (>80)
Child Has Never Attended School
Child Has Attended School



3.4 Differences by socioeconomic status

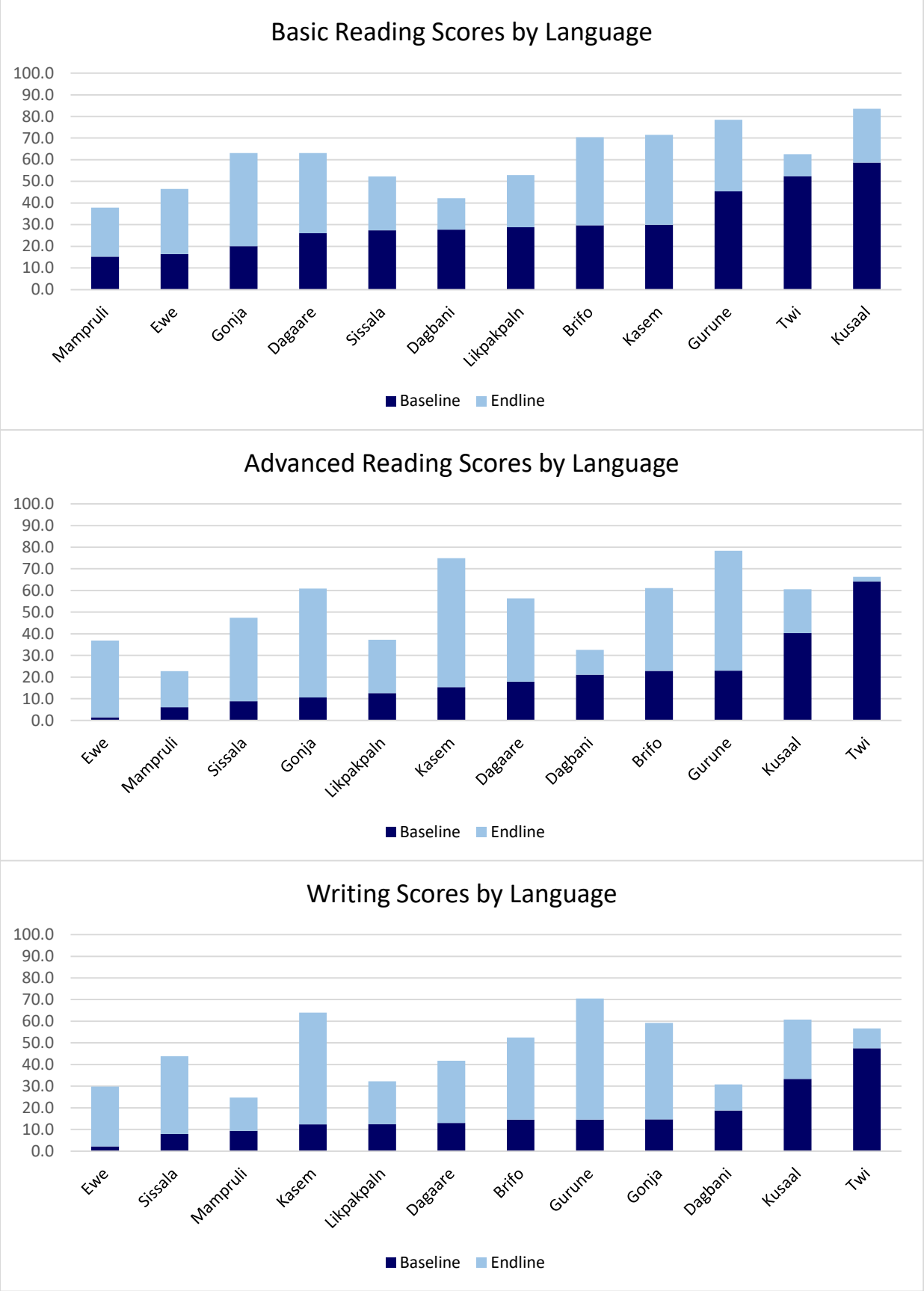
In the baseline results, it was shown that student performance did not vary significantly by socioeconomic status. This was explained by the assertion that there was too little variation in socioeconomic status in the CBE population to have any real predictive power. Since the wealth index variables were not collected again at endline, the variation remained what it was in the baseline report. Accordingly, there was a similar lack of relationship between socioeconomic status and endline scores (as well as improvement). Therefore, no graphs have been included in this section.

3.5 Differences by language

As previously noted, students were assessed in one of 12 Ghanaian languages, corresponding to the language in which they were learning within their CBE centre.

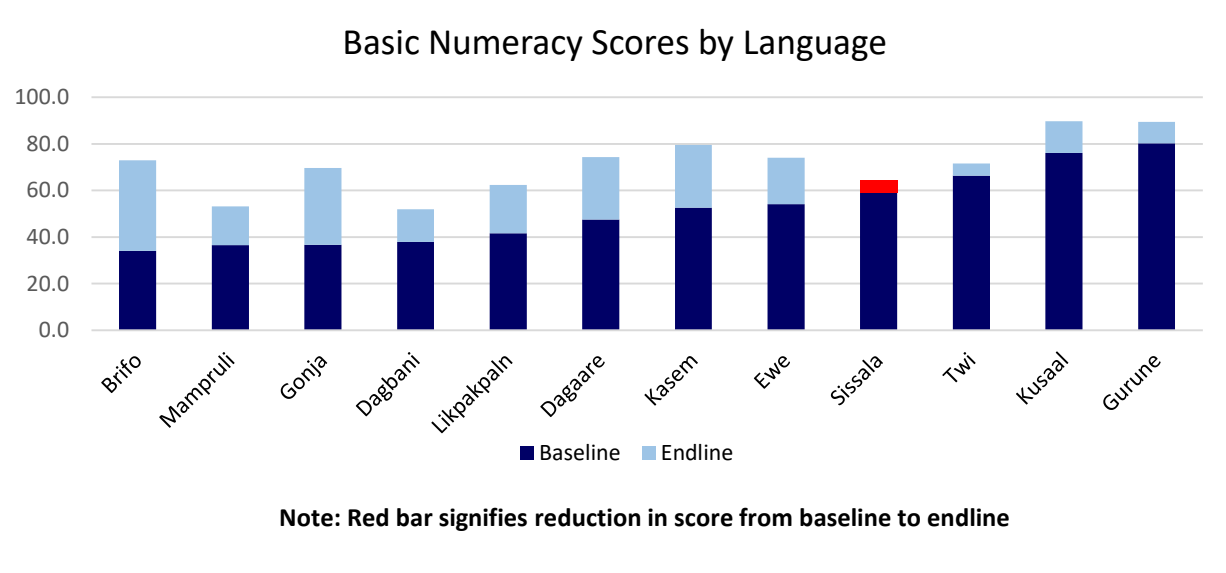
Since it is complicated to review output for changes in performance for 12 languages on five competencies across four categories, the outputs in this section focus on the composite scores themselves (as opposed to the student profile categories). Each figure provides an overview of the baseline scores (in dark blue) with the growth in performance by endline (in light blue). All figures are sorted from lowest to highest score on the baseline assessment. The results from the three literacy components (basic reading, advanced reading, and writing) are presented in ***Figure 8.*** The most obvious finding from this figure is that the relative performance of students in a given language was not consistent over time. In other words, students speaking some languages that charted low performance at baseline improved much more than their counterparts speaking languages with higher performance – in which cases they either narrowed or even reversed the gaps. For example, while Twi students were the highest performers in advanced reading and writing (and second highest in basic reading) at baseline, they were outperformed by pupils using four other languages in basic reading and writing (and two other languages in advanced reading) at endline. At the other end of the spectrum, Kasem students were the fourth-lowest performers in writing (and middle performers in advanced reading) at baseline but improved to the second-highest spot in both competencies by endline. Gonja and Gurune students also showed impressive gains in all literacy competencies, putting them in the top group of performers across the board. Some of this may be explained by the language make-up of particular CBE centres, districts, and/or regions, as multi-language contexts are likely more difficult to implement in homogeneous settings. This issue is explored later at the end of this section.

Figure 8. Literacy composite baseline and endline scores, by language

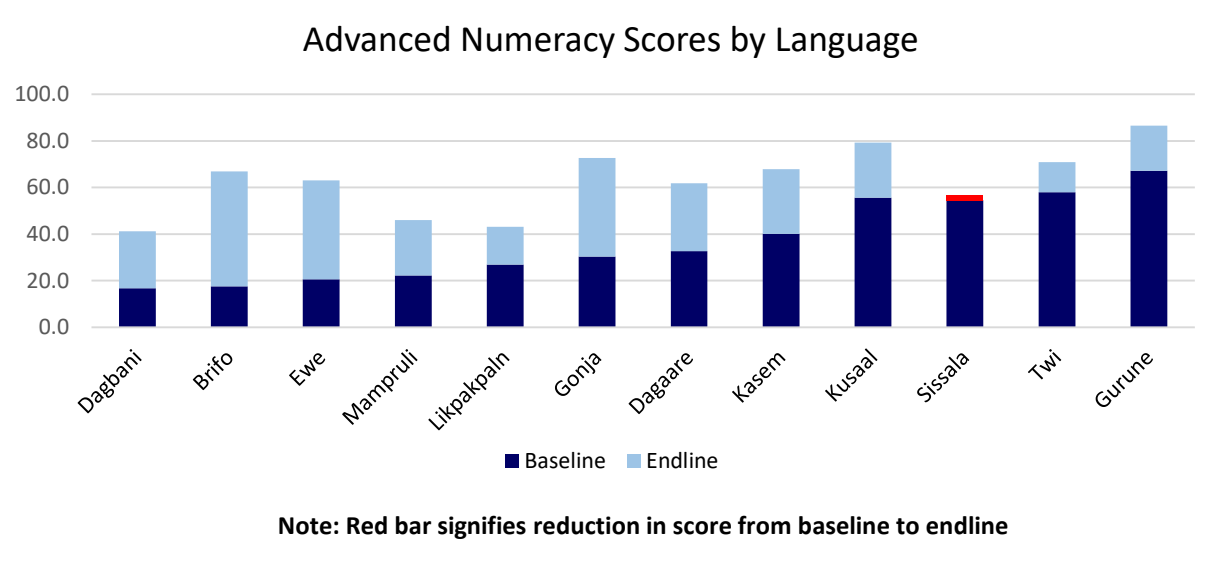


The results were a little more consistent for numeracy, but several lower-performing languages still showed large improvements and higher-performing languages did not remain at the top ***(Figure 9).*** Brifo students moved up to the middle of the performance spectrum by endline, having started as the lowest and second-to-lowest performers at baseline in basic and advanced numeracy, respectively. Kusaal and Gurune students both began and ended near the top of the language groups. Gonja students once again showed impressive gains from baseline to endline (as was the case with literacy, above). The most interesting (and confusing) finding is that Sissala students scored in the top third for both basic and advanced numeracy at baseline but then had lower scores on the endline assessment (i.e. their performance went down over the course of the nine-month programme in both numeracy competencies).

Figure 9. Numeracy composite baseline and endline scores, by language



Note: Red bar signifies reduction in score from baseline to endline



Note: Red bar signifies reduction in score from baseline to endline

3.5.1 Student language compared with centre language of instruction

One additional factor that could influence the impact of language on student performance is the match between a student's home language and his/her CBE centre's language of instruction. Overall, this

match rate was very high. Specifically, 92% of students claimed that they spoke the same language at home as that which was used in the CBE centre. Furthermore, 84 of the centres had 100% language match and only 17 centres had less than 90% match. There were, however, some exceptions. In the most extreme case, there were 3 centres in the Northern region with zero percent language match, 1 centre in Ashanti with 12.5% match, and 2 additional centres in the Northern region with less than 50% match. While there are no significant differences in performance gains (on any of the competencies) for those students who do and do not speak the language of instruction, this is driven to some extent by the fact that the sample of students/centres without a match is very small. Therefore, since the main focus of these centres is to provide language of instruction in the home language of students, it will be important to follow up with these particular, problematic centres to learn more about why such a high mismatch occurs (particularly in those centres where no students speak the LOI).

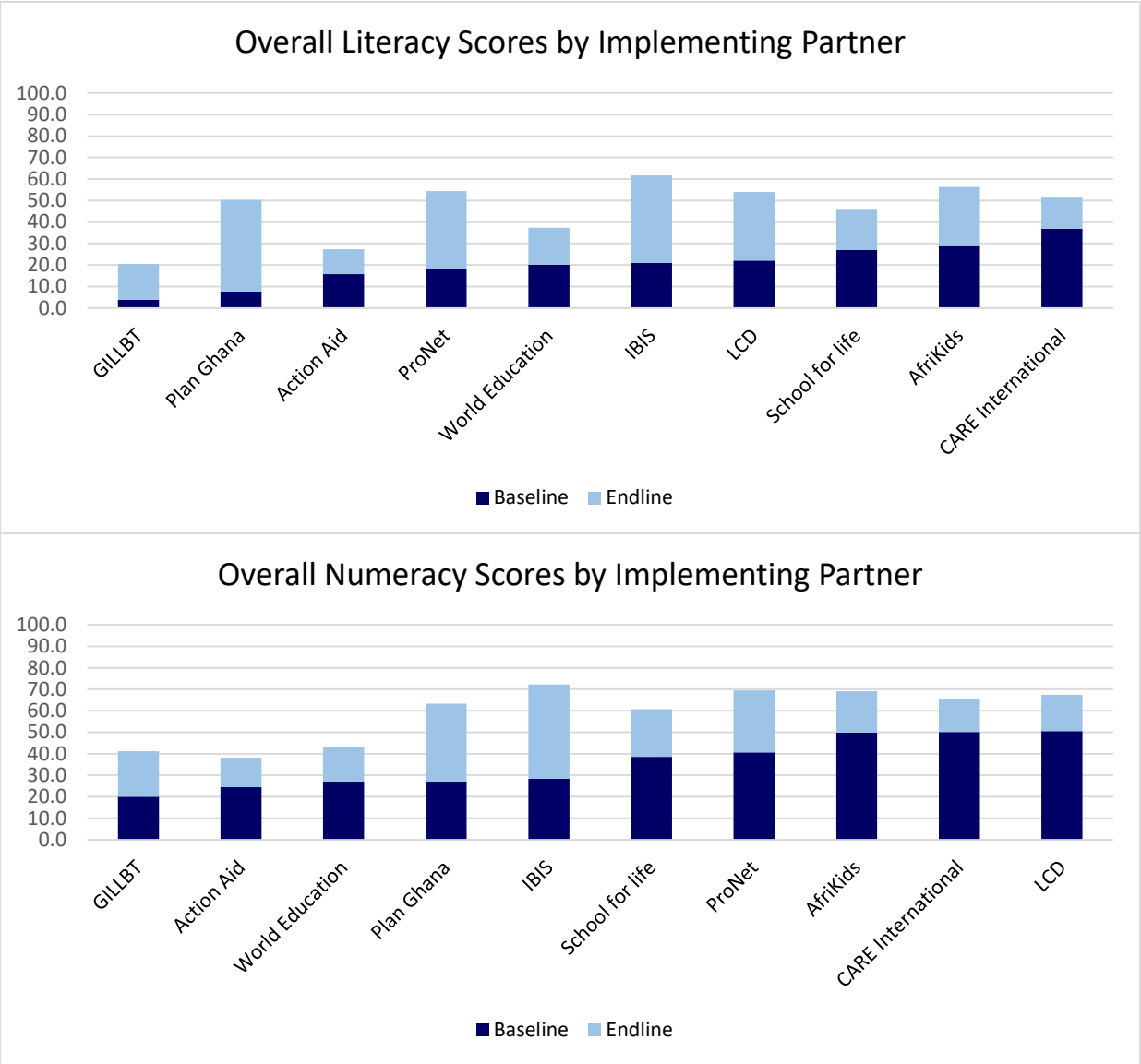
3.6 Differences by implementing partner and region

CBE centres in this study are operated by 10 different implementing partners. These implementing partners work in particular regions or districts of Ghana and teach students in particular languages appropriate for that region. Accordingly, there is a very high correlation among implementing partner, region, and languages – and it is very difficult to disentangle the three. Qualitative research (outside the scope of this study) could provide deeper insights into the performance of specific implementing partners.

Comparing student performance across implementing partners and regions led to some inherent redundancies with the language groups presented above (due to the high rate of overlap between language, IP, and region), but because these values are of interest to CBE evaluators and policy makers, results are provided below. The results are presented for overall literacy and numeracy scores only (as opposed to separately for each competency). It should be noted that these analyses must be seen not as evaluations of IP performance but simply as measures of different endline growth estimates in performance across IPs (which were confounded with a number of factors, including language). Overall literacy and numeracy scores (from baseline to endline) as disaggregated by implementing partner, are displayed in ***Figure 10.***

This figure displays baseline scores in dark blue with endline gains in light blue. The top of each column is therefore the total endline score. Both the literacy and numeracy figures are sorted from smallest to largest baseline scores. While all implementing partners improved significantly from baseline to endline, a few implementing partners appeared to improve more than the rest. For both literacy and numeracy, Plan Ghana, ProNet, and IBIS showed larger gains than their counterparts. IBIS moved from the middle of the pack at baseline to become the top endline performer in literacy and numeracy.

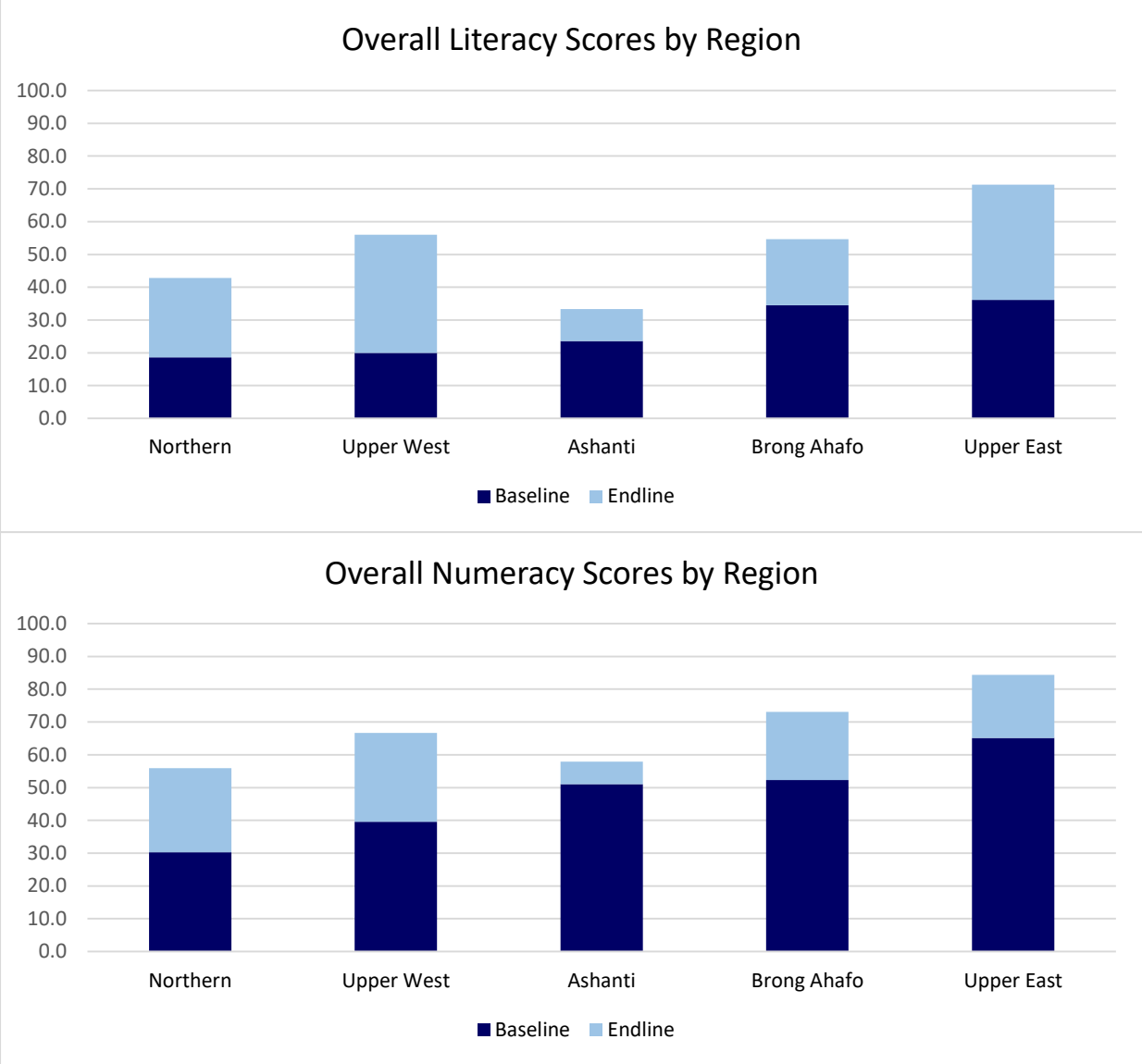
Figure 10. Overall literacy and numeracy score gains, by implementing partner



Note. These values are not evaluations of IP performance; they are measures of different endline growth estimates in performance across IPs, without taking into account various confounding factors.

The results for overall literacy and numeracy score gains (baseline to endline), disaggregated by region, are displayed in ***Figure 11.*** As was the case for language and implementing partner, all regions improved from baseline to endline. What is clear from this figure, however, is that students in the Ashanti region made the smallest overall gains. While Upper West students may have produced the largest overall gains in literacy and numeracy, there was only a minimal difference between their gains and those of the remaining three regions.

Figure 11. Overall literacy and numeracy score gains, by region



4. Understanding sources of variation in literacy and numeracy test score gains

4.1 Ordinary least squares (OLS) regression modelling

While Chapter 3 examined variations in literacy and numeracy scores (including gains) for a range of factors independently, for this chapter, all variables of interest were included in a single model to allow examination of the relative impact of each.

Accordingly, the results from five OLS regressions are displayed in ***Table 10***. The five models in the table each represent a regression with the same predictor (independent) variables but different outcome (dependent) variables. Specifically, the models in Table 10 used the gain scores for basic reading, advanced reading, writing, basic numeracy, and advanced numeracy. All dependent variables were calculated by subtracting baseline scores from endline scores (for a given composite category) and include the baseline score as an independent variable in order to account for the starting point.

Table 10. Linear regression models predicting gain scores

Explanatory variables	Model 1 Basic reading	Model 2 Advanced reading	Model 3 Writing	Model 4 Basic numeracy	Model 5 Advanced numeracy
Age	1.15**	1.94***	1.44**	1.19**	1.99***
Facilitator Assistance	3.97**	1.62	1.49	3.77**	3.41**
Group Work	1.96	4.03**	2.94*	2.44*	4.52**
Pupil home language equals language of instruction	12.92**	10.97	11.27	8.46*	3.71
Female	-3.09	-3.67*	-2.79	-1.56	-0.31
Language					
Brifo	17.89***	11.82	14.87*	14.13*	10.45
Dagaare	18.62**	18.01*	11.54	15.82*	10.24
Dagbani	-3.96	-9.51	-5.09	-3.44	-8.10
Ewe	2.81	-0.30	-0.10	7.66	7.71
Gonja	26.97***	29.78**	32.63***	18.87**	20.86**
Kasem	23.48***	32.23***	33.89***	15.72**	9.24*
Kusaal	16.60***	3.72	10.97	11.92*	7.14
Likpakpaln	1.52	-5.42	-3.71	1.32	-15.29**
Mampruli	-0.16	-12.48	-5.75	0.14	-3.60
Sissala	2.70	2.93	5.83	-11.93	-16.30
Twi^	0.00	0.00	0.00	0.00	0.00
Gurune	13.64**	25.67**	27.59**	7.19	6.24
Number of people who live in the same house	0.09	0.00	-0.10	0.05	0.06
Does any activity at home that involves reading or writing	1.20	2.46	3.67	1.26	4.24
Has ever attended school?	1.12	2.46	2.77	0.77	-1.63
Number of days missed in the last 5 school days	0.12	-0.50	-1.06	-0.61	-0.56
Does any work outside of home	0.28	-0.62	-0.71	-0.47	-0.40
Wealth index					
Low^	0.00	0.00	0.00	0.00	0.00
Mid-low	-6.05**	-5.56*	-4.14	-3.37	-2.52
Mid-high	-7.75**	-4.21	-3.88	-4.74*	-4.34
High	-5.17*	-1.36	-1.15	-4.27	-4.48
Baseline score	-0.46***	–0.56***	–0.43***	–0.56***	–0.60***
Constant	-0.74	-4.44	-5.07	8.92	3.35

Explanatory variables	Model 1 Basic reading	Model 2 Advanced reading	Model 3 Writing	Model 4 Basic numeracy	Model 5 Advanced numeracy
N	1877	1877	1875	1860	1860

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$. ^ = reference group.

Aside from the wealth index and language variables, the interpretation of all values in Table 10 is that a one-unit increase in the variable in the first column was associated with a gain score increase or decrease equivalent to the coefficient displayed for each model (positive means increase; negative means decrease). For example, a one-unit increase in age (i.e. being one year older) was associated with a gain in the basic reading score of 1.15 points (from baseline to endline). For advanced reading, the associated gain score for an extra year of age was 1.94 points – meaning that age had a larger impact on improvements in advanced reading scores than it did on improvements in basic reading scores. Additionally, it should be noted that these coefficients are interpreted as the increase 'holding all else constant', meaning that the value shown is the independent increase associated with age, after accounting for all other variables in the model.

For the wealth index and language variables, each coefficient is interpreted as the gain score impact relative to the reference group. For the wealth index, the reference group was the lowest quartile (meaning that all other quartiles were compared directly to the low wealth group). Accordingly, in some of the models it was found that students with the lowest wealth index scores produced significantly higher gains than those students from higher wealth quartiles. This was particularly apparent in the basic reading gain score model. For language, the reference group was Twi (e.g. Brifo was associated with a gain score in basic reading of 17.89 points larger than Twi, even after accounting for baseline scores).

The relative impact of each factor included in Table 10 was generally similar across models. An additional year of age was associated with a gain score increase between 1.15 and 1.99 points, depending on the component. This means that older students improved more than younger students. Being female was negatively associated with gain scores (i.e. girls saw lower gains) but only in the advanced reading model was the coefficients statistically significant.

There were three factors from the student questionnaire related to the CBE programme that were found to significantly impact gain scores. The first was a question on facilitator assistance, specifically "When you did not understand, did your teacher help you to understand?". Students answered on a four-point scale: 0="no"; 1="only sometimes"; 2="yes, most of the time"; 3="yes, all the time". Notably, there was a 3-4 point increase in gain scores for basic reading, basic numeracy, and advanced numeracy for each additional point up the four-point scale. In other words, the more often a facilitator helped a student understand when the work was difficult, the larger the increases were for those students.

Additionally, participating in group work was associated with greater gain scores than not participating in group work for all models except basic reading. This variable was on the same 4-point scale as facilitator assistance and the impact was similarly large, ranging from approximately 2.5-4.5 points.

Third, in terms of language, coefficients were relatively large for language match (i.e. pupil home language equals language of instruction). Even after accounting for the languages themselves, it was found that having a language match for a student was associated with a nearly 13 point gain in basic reading (as compared with those students without a match). The result was also statistically significant for the basic numeracy model, with a language match providing an 8.5 point gain over the improvement of their non-matched peers.

As for specific language differences, Twi-speaking students scored higher on both literacy and numeracy subtasks than students in other languages at baseline. However, Table 10 shows that some of the language groups that were struggling at baseline made large strides by endline. For example, Brifo and Kasem students showed significantly larger gain scores than their Twi

counterparts on all five competencies. Gonja, Dagaare, Kusaal, and Gurune students all improved more than Twi students for at least three of the competencies. Sissala and Dagbani were the only two languages to have students with significantly smaller gains than Twi students in any of the models. In terms of the overall reading score (not shown in this table), Twi students had the highest average score at baseline but they were overtaken by Gurune, Kusaal, and Kasem students (and matched by Gonja and Brifo students) by endline. For the overall maths score (also not shown in this table), Brifo students had the lowest baseline average but by endline, they had improved to the point where only three languages were outperforming them (Kasem, Kusaal, and Gurune students).

Interestingly, none of the five background items from the child questionnaire included in these models (i.e. family size, reading/writing activities at home, prior school attendance, student absenteeism, and work outside the home), were significant. These insignificant findings point to the fact that other variables in the model account for the changes in performance that are sometimes attributed to these factors in independent models without controls. Similar findings were found in a regression using implementing partners in the model in place of languages. Output from those models can be found in ***Annex 7.***

4.2 Predicting dropout

Results from basic sample attrition analyses were presented in Chapter 2 of this report. As noted in that section, however, the current data only provide information on sample attrition and do not allow for any investigations into dropout. The next round of data colleciton (i.e. the tracker study for Cycle 4 students) will be used to obtain infomration on completion rates as well as reasons for dropout (among those students who were unable to finish the CBE programme). Accordingly, dropout analyses will be conducted using Cycle 4 endline performance data once information on dropout is collected.

5. Conclusions and Recommendations

Overall, the majority of students (80–85%) in the CBE programme significantly improved their literacy and numeracy skills during the nine-month cycle. Furthermore, there were large increases in the mean performance scores for students across all 15 of the subtasks in the literacy and numeracy assessments. While students continued to underperform on the literacy subtasks (as compared with numeracy), the average gains in literacy were higher. On average, mean literacy scores increased at a nearly consistent 25 percentage points. The average increase in numeracy scores showed more variation, ranging from 17 to 29 percentage point gains. Results also showed large improvements across all proficiency score categories from baseline to endline, in both literacy and numeracy. These results all point to strong learning gains from the programme. This is further evidence that the programme was effective in ensuring that students were better academically prepared to transition into the formal school sector than when they began the programme.

Investigations of learning gains by subgroups showed that male students were more likely to become proficient in literacy and numeracy than their female counterparts. This created a small separation in the performance that was not found at baseline. While no significant differences were found in endline performance or gains based on wealth, differences were found in terms of prior school attendance. Counter to the emerging gap in performance by males and females, students with prior school attendance outperformed their unschooled counterparts at both baseline and endline, effectively narrowing the gap that existed at baseline. While it is unclear why boys outpaced girls over the course of the programme, it was very promising to see that students without prior school experience were able to make significant gains toward equity with those who had been to school. Ultimately, this study shows that the CBE programme may be particularly well suited for students without prior school experience (especially before they enter the formal school system) – though the large gains for students with prior experience show that the programme should continue to focus its efforts on both types of out of school children. In the next phase of the study – when CBE students are tracked through a year of traditional primary schooling – particular attention should be paid to the impact of

prior school attendance on student performance and retention. This finding is not to say that we shouldn't focus on those that dropped out but simply that those who had never been to school started at a lower point and started catching up to those that had. It is important that the programme continues with both groups.

Although performance gains were found for nearly all languages, implementing partners and regions, the relative performance was not consistent over time. These data are not sufficiently nuanced to investigate why particular language groups and IPs improved more than others, but these findings should definitely be an area for further research. A qualitative study (focused specifically on high and low achieving centres) would be invaluable in order to learn from the languages and implementing partners that showed greater gains, and to improve upon those with smaller gains. Because the data made it possible to determine which centres had the smallest and largest gains, a case study approach could examine what the specific centres were doing to improve performance with greater and lesser degrees of success.

When we examined variables jointly in multivariate regression models (as opposed to independently in the previous analyses), it became evident that a few variables had significant predictive power on student performance gains. For example, student age was associated with gains (older students showed larger gains) for all competencies. Additionally, facilitators helping students when they had a difficult time understanding and engaging in group work were both found to significantly increase gains across a number of competencies. Accordingly, ensuring that facilitators are regularly providing constructive feedback to students and training them to incorporate productive group work activities should both remain as key programme components in future cycles of CBE. Lastly, in terms of language match, these models provided evidence that having a match between a student's language and the centre's language of instruction is particularly important for the most basic skills (i.e. basic reading and basic numeracy). This is not to say that language match isn't important for the more advanced skills but simply that there may not have been enough variation in this variable (with such high rates of match) to produce significant coefficients. Either way, CBE centres should undoubtedly continue to focus on providing instruction in children's home languages and follow-up work should be conducted to learn more about why high proportions of match are not always possible and what steps should be taken to ensure that all students have the best opportunity possible in multi-language centres.

Ultimately, it is clear that students in CBE centres significantly improved in their literacy and numeracy scores during their nine-month programme. While it appears that these students are now more prepared for the formal school system than they were nine months ago, comparing their performance to that of traditional public school students in the next round of this study will shed further light on the ultimate impact of the CBE programme on school readiness.

Annex 1. Child Background Questionnaire (Baseline)

COMPLEMENTARY BASIC EDUCATION (CBE) PROGRAMME IN GHANA
CHILD BACKGROUND QUESTIONNAIRE

Instructions:

- Administer this questionnaire to every child BEFORE they take the tests.
- Call each child from the classroom, administer the questionnaire and let them return to their class.
- After every child in your sample has completed this questionnaire, call each one in the same order to do the tests [this is to ensure that they are not overburdened on the day of the test]. Make sure that you interview each child at a time, and that they return to their class before calling the next child on your list.

Verbal Consent

- Explain to the child in their local language the reason for asking further information. Explain SLOWLY that you want to know a little more about their previous schooling experience [before they enrolled on the CBE programme], and some information about their household.
- ASK EACH CHILD IF THEY UNDERSTAND WHAT YOU ARE ABOUT TO ASK THEM AND ARE HAPPY TO ANSWER.

1. Information

Date of interview: (DD/MM/YY)	[pre-programmed in DDG]
Name of interviewer:	[code pre-programmed in DDG]
Name of IP provider:	[pre-programmed in DDG, select from list]
Name of the CBE Centre:	[pre-programmed in DDG, select from list]
Language:	[pre-programmed in DDG, select from list]
Region:	[pre-programmed in DDG, select from list]
District:	[pre-programmed in DDG, select from list]
Child ID number:	[we need a unique id number to make sure that we track children over time, and that we can use this number to link the different datasets]
Name of Child:	
Age of Child	Years
Gender: Boy or Girl	01=Boy; 02=Girl;

2. Family Status:

2.1	Who looks after you most of the time at home? Enter number:	
01=Mother; 02=Father; 03=Grandparent; 04=Aunt / Uncle or other relative; 05=Foster Parent		
2.2	How many people live in the same house as you live in?	
2.3	How many brothers do you have?	
2.4	How many of your brothers go to school?	
2.5	How many of your brothers have attended this type of school you are attending?	
2.6	How many sisters do you have?	
2.7	How many of your sisters go to school?	
2.8	How many of your sisters have attended this type of school you are attending?	
2.9	Do you have books or any reading material at home to read? 01=yes, 02=No	
2.10	Do you do any activity at home that involves reading or writing 01=yes, 02=No	

3. Household economic situation

3.1	Does anyone in your house own a mobile phone? 01=No; 02=YES; 99=NK	
3.2	Does anyone in your household own a bicycle? 01=No; 02=YES; 99=NK.	
3.3	Does anyone in your household own a motor bike? 01=No; 02=YES; 99=NK.	
3.4	Does your family own a radio? 01=yes, 02=No; 99=NK	
3.5	Does your family own a TV? 01=yes, 02=No; 99=NK	
3.6	During the night, which of the following do you mainly use to give you light? 01=light bulb; 02=kerosene lamps; 03=candles, 04=touch light; 05=firewood; 06=solar lamps	
3.7	How often does your family have enough food? 01=everyday; 02=some days we go hungry; 03=most days we go hungry	
3.8	Compared to other families in your village/town do you think your family has: 01=more money; 02=the same money; 03=less money	

4. School History

Please ask the following questions about the child experience

	Have you ever been to school? 01=yes; 02=no	
4.2	If yes, how many years did you go to school?	
4.3	In which class did you stop school? 01=P1;02=P2; 03=P3; 04=P4; 05=P5; 06=P6; 07=JHS	
4.4	What was the reason for stopping?	

	01=sickness; 02=lack of money; 03=distance to school; 04=help at home 05=getting married/have baby, 06=work outside home, 07=repeated too many times; 08=other_____	
4.5	Did you ever repeat a class? 01=No; 02=YES; 99=NK	

5. Current School

5.1	Apart from this school do you go to another school? 01=yes, 02=no	
5.2	Think about the last five school days, how many days have you missed school? 01=1 day, 02= 2 days; 03=3 days; 04=4 days; 05=5 days	

6. Work

6.1	Do you do any work at home? 01=yes; 02=no	
6.2	What kind of work do you MOSTLY do? 01=help with fetching water; 02=help with cooking; 03=help on the farm; 04=help with cleaning the house/compound; 05=caring for siblings, 06=other_____	
6.3	Do you do any work outside of home? 01=yes; 02=no	
6.4	What type of work do you MOSTLY do outside the home? 01=work on the farm; 02=help keep livestock; 03=petty trading	
6.5	If you work outside the home, do you get paid for the work you do? 00=No; 01=YES; 99=NK	

7. Child Opinions

Please read these statements to the child in their local language and enter the appropriate response. Please ask the child to refer to their last school experience. Ask these questions to those who have ever been to school

Instruction: I am going to read out some sentences, listen to each of them carefully and tell me whether you strongly agree; agree; disagree or strongly disagree. These statements refer to your last school experience and NOT your current CBE school experience

Code: 01= all the time; 02=most of the time, 03= sometimes; 04=never

READ THIS EXAMPLE. When my mother cooks food (appropriate food in the area for example fufu), I enjoy it. (All the time, most of the time, sometimes, never). The children should pick one.

		Response
Being at School		
7.1	I felt happy when I was at school	
7.2	I was often tired when I was at school	
7.3	I did not like going to school	
7.4	School lessons were not interesting to me	
7.5	The language the teacher used was easy for me to understand	

		Response
7.6	The language the teacher used was my own language.	
7.7	I felt safe at school	
7.8	I was often hungry	
7.9	I was often beaten by teacher	
7.10	I was often beaten/mocked by my friends	
Learning		
7.11	I found most lessons easy when I was at school	
7.12	I tried hard to learn my lessons	
7.13	I was very good at mathematics at school	
7.14	I found learning language difficult	
7.15	I found learning easier when I was taught MOSTLY in my mother tongue	
7.16	When I did not understand my lesson, I asked my teacher for help	
Sources of support at home		
7.17	When I did not understand things at school I asked my mother or female adult	
7.18	When I did not understand things at school I asked my father or male adult	
7.19	I asked my brothers or sisters to help with my school work	

THANK YOU FOR ANSWERING MY QUESTIONS.

LATER I WILL CALL YOU TO DO A SIMPLE TEST.

YOU CAN NOW GO BACK TO YOUR CLASS.

Annex 2. Sample Literacy Assessment: Twi

Complementary Basic Education (CBE) Programme

Tools for Assessment of Learners' Literacy and Numeracy Skills

Test and Item Specifications

Cover Page

[Asante Twi]

Demographic Information on the Learner *(Complete)*

Name of Learner	_____	Age	_____	**Sex**	_____
District	_____		**Community**	_____	
Language Used at Home (First Language)			_____		
Schooling Status (*Tick*)					
i	Never been to school		YES	NO	
Ii	Drop out of school		YES	NO	
If drop out, at what stage/grade level did learner drop out of school? (*Tick*)					
I	Lower primary				
Ii	Upper primary				
Iii	Junior high school				
Enumerator Information (*Complete*)					
Enumerator ID	Name of Enumerator		Phone contact		
01					
02					

Atwerɛ ne Akenkan Mpɛnsɛmpɛnsɛmu Kwan

Saa atwerɛ ne akenkan mpɛnsɛmpɛnsɛmu nwoma yi fa dwumadie atitire 6 a ɛdidi soɔ yi ho:

1. Atwerɛdeɛ/nnyegyeeɛ yikyerɛ:

   Merebɛtwerɛ atwerɛdeɛ 50 agu ɛpono no so. Mepɛ sɛ mo mu biara bɛbobɔ atwerɛdeɛ no a ɛwɔ ɛpono no so mmaako mmaako ama matie. Sɛ mowie a, metwa mo dwumadie no.

1	2	3	4	5	6	7	8	9	10		Nkasaho
a	f	n	m	u	y	s	w	g	t	(10)	
ky	p	b	hw	r	l	d	e	h	k	(20)	
gy	ny	a	p	t	K	f	nw	hy	s	(30)	
d	o	g	i	ky	y	e	m	n	ny	(40)	
ɛ	ɔ	b	u	r	l	hw	h	hy	f	(50)	

2. Nnyegyeeɛ ho nimdeɛ –

   Merebɛma mo nnyegyeeɛ 10 wɔ Asante Twi mu. Mo mu biara bɛma me nsɛmfua wɔ Akan kasa mu a saa nnyegyeeɛ yi na ɛhyɛ aseɛ wɔ **sima num** ntam.

Nnyegyeeɛ	Nsɛmfua nhwɛsoɔ	Deɛ edi mu	Deɛ enni mu	Deɛ mmuaeɛ amma
/s/	sane			
/f/				
/t/				
/k/				
/h/				
/p/				
/b/				
/d/				
/m/				
/w/				

3. Nimdeɛ wo nsɛmfua hunu mu (Nsɛmfua a wɔnim dada)] - Merebɛtwerɛ nsɛmfua 20 agu ɛpono no so ɛno nti monhwɛ no yie. Mo mu biara nkenkan no wɔ sima mmienu ntam.

Nsɛmfua a wɔnim dada akenkan dwumadie

papa	owura	nipa	wie	tam
pia	bobo	pusu	kyerɛ	him
bɔtɔ	pete	koto	pɛtɛ	sika
hyɛ	kete	kuta	dwetɛ	tutu

4. Akenkan ne Nteaseɛ dwumadie –

   Montie abasɛm tiawa a merebɛkenkan no yie. Sɛ mewie a, mɛbisa mo nsɛm afa abasɛm no ho.

   Kofi Opoku **y**ɛ abarimaa. Ɔyɛ obi a ɔwɔ ahobrɛaseɛ ɛna obu adeɛ. Kofi Opoku kɔ sukuu wɔ Daakye Nti D/A Mfitiaseɛ sukuu. Ɔgyina klase 2. Ɔyɛ obi a ɔnim adeɛ yie. Ɛnam Kofi Opoku suban pa ne ne mmɔdemmɔ nti, atikyafoɔ a wɔwɔ sukuu mu hɔ no nyinaa pɛ n'asɛm.

Nsɛmmisa

1. Ɛhenefa na Kofi Opoku kɔ sukuu?_____

   _____

2. Klase bɛn na Kofi Opoku gyina?_____

   _____

3. Ɛdeɛn na osuani bɛyɛ a ne tikya mpɛ n'asɛm?_____

   _____

4. Sɛ obi da suban pa adi na ɔbɔ ne ho mmɔden a mfasoɔ bɛn na ɛwɔ so?_____

   _____

5. Akenkan anonkwokwasoɔ: _____

   _____

   Monhwɛ abasɛm a ɛwɔ ha yi. Mo mu biara rebɛkenkan no den akyerɛ me wɔ **sima baako** ntam.

   Dabiara anɔpa sɛ meda nyane a, mehohoro m'anim, na matwitwiri ne se. Me maame dware me wie a, na masera, manunu me tiri, na mabɔ pawda. Mewie a, na madidi. Sɛ medidi wie a, na mafa me sukuu baage na makɔ sukuu. Sɛ mepɔn sukuu awia a, na maba fie.

6. Nsɛmfuatwerɛ–
   Mo nyinaa ntie no yie. Merebɛbobɔ nsɛmfua num na mo mu biara atwerɛ no yie agu ne krataa mu.

   i. _____
   ii. _____
   iii. _____
   iv. _____
   v. _____

7. Nsɛmfua num no na ɛdidi soɔ yi.

   i. Kɔtɔ
   ii. Poma
   iii. Abofra
   iv. Sukuu
   v. Owia

8. Bɔsramuka atwerɛ/Ɔkasamuyɛ–

   Obiara ntwerɛ kasamufa tiawa a ɛtɔ asom **mmienu** a ɛfa aduane a ɔpɛ ho.

   i. _____
   ii. _____

Annex 3. Sample Numeracy Assessment

1. Identification of Single Digit Numbers

This is a timed test of number recognition. Before starting the timed test help the learner to feel relaxed. Let them see the chart and explain carefully what they have to do. They should say each number across the line and you score how many are correct within one minute.

1	2	4	6	8	3	5	7	9	4
3	5	1	4	6	8	7	2	3	9
4	2	9	2	3	6	5	4	7	1
2	4	6	5	3	7	1	8	3	9
4	3	7	1	6	8	2	9	4	5

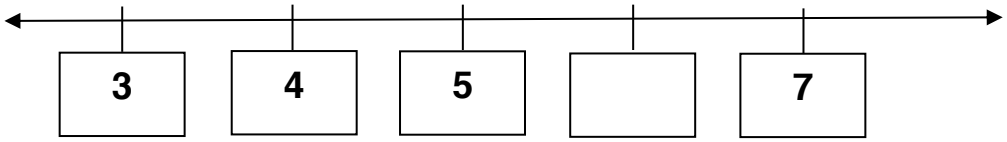
2. Identification of Two Digit Numbers

Follow the same procedure as before. Record the total number recognised by the learner in one minute.

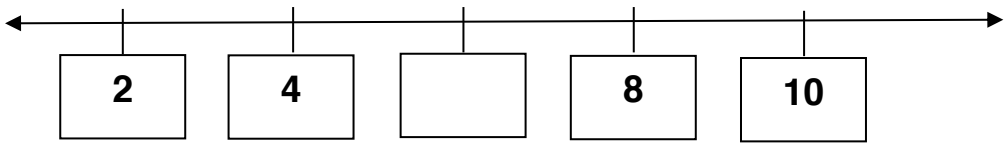
I	Ii	Iii	iv	V	Vi	vii	viii	Ix	X	
10	14	18	20	22	24	28	34	43	50	(10)
11	16	23	33	25	15	51	30	36	42	(20)
19	12	26	17	46	64	35	13	31	32	(30)
45	54	55	86	68	78	89	90	95	99	(40)

3. Missing Numbers

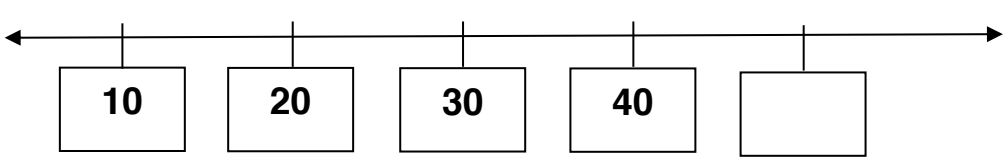
(1)



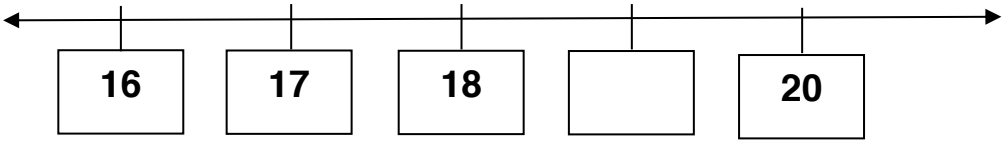
(2)



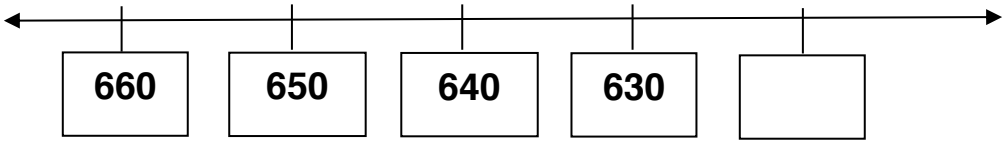
(3)



(4)



(5)



4. Problem solving

This should be given to the whole class as a written test. Explain to the learners that they should try to solve as many of the questions as possible. Explain that they should work independently and should not confer or copy other learners. Allow 40 minutes for completion.

Addition	**Subtraction**
1. 5 + 6 = ☐ **2. 8 + ☐ = 10** **3. Asana has 8 mangoes. Adam gave her two more mangoes. How many mangoes did Asana have altogether?**	**1. 8 – 6 = ☐** **2. ☐ – 5 = 4** **3. Nadama bought 10 eggs and 4 got broken. How many of the eggs are left unbroken?**
1. 25 + 14 = ☐ **2.** 37 + 24 = ☐ **3.** 54 + 23 = ☐ **4.** 138 + 247 = ☐	**1.** 35 – 14 = ☐ **2.** 43 – 24 = ☐ **3.** 82 – 54 = ☐ **4.** 362 – 145 = ☐
Multiplication	**Division**
1. 10 X 3 = ☐ **2. ☐ X 4 = 12** **3. 6 X 5 = ☐**	**1. 10 ÷ 2 = ☐** **2. 12 ÷ ☐ = 4** **3. Mr Karimu gave 12 books to his three children to share equally. How many books will each child get?_____**

Annex 4. Report on the Learning Assessment Training

The training began with sharing of experiences from the baseline study. The general lessons learnt included the following:

1. Access to most of the communities was difficult and they were worried that when the rains set in, access was going to be more difficult for the endline study.
2. They were surprised at the conditions under which the children were studying and commended the efforts of the donor agency and the government in ensuring that the CBE was being organised.
3. Notwithstanding the above, they were also impressed about the eagerness of some of the pupils to study and the dedication of the facilitators.
4. In general, they found the female facilitators more committed than the male counterparts.
5. They were happy about the facilitating support received from the IPs in conducting the study.
6. They appreciated the use of the tablets and the mutual benefit received from the WhatsApp platform for internal communications.

In the training itself, the participatory question-and-answer approach and the use of role play methods were adopted to go through the questionnaire. The major question that was raised was that the format of some of the numeracy questions had been altered from the baseline questionnaire. This is an example:

Baseline: 2 + 6 = 8☐

End line: 2 + ☐ 8

It was explained that the change was not significant and if any child had memorised the baseline question, this would be the test to know whether he or she had learnt anything new after eight months of CBE classes.

Pilot study

In this endline training, a new set of qualitative data were to be collected which were not part of the baseline survey. These were Child Pedagogy and Teacher Pedagogy Observations. After the classroom training, a pilot study was organized in one of the CBE centres in the Tolon district. Three sets of data were to be collected:

1. Observation of the facilitator's approach to teaching
2. An interview with the facilitator on his or her background and perception of the CBE
3. Observation of the attention/participation of eight selected pupils.

There was also an audio recording of the proceedings and the interview with the facilitator. After the community entry protocols, three enumerators were selected to undertake the data collection while the rest interacted with the community members.

At the end of the classes, the team orally submitted this report on the facilitator:

1. The facilitator's observation was good, using all participatory methods, gender balanced in questioning, etc.
2. He encouraged students to speak and asked for clapping for good answers and did not shout at wrong answers.

3. He invited students to the blackboard to identify words they had pronounced, etc.

Two enumerators reported on the facilitator and the summary was quite good and encouraging about his performance. But something else was observed: The lesson went too smoothly and too lively, giving the impression that he was not teaching a new topic. Also very few pupils got wrong answers; it appeared the topic had been rehearsed for the observers' consumption. The recorded interview was played back and the enumerators were advised to ask the facilitator about this issue if they got such an impression when they eventually went to the field.

In reporting on the pre-selected 8 pupils, the following was found:

1. The enumerator could follow and report on only 3 of the 8 sampled pupils (we had opportunity for only one lesson)
2. She was able to report on their participation and responses in class accurately
3. However, the events she reported on were not chronologically followed
4. Interestingly, she was trapped by the performance of the facilitator and reported on him as well instead of concentrating on the pupils.

It was further reported that the facilitator declined the interview on the excuse that he wanted to start the second lesson immediately and would also attend the Ramadan prayers after 5pm. This reinforced an earlier decision in class that facilitators' interviews should be held on a second day before the start of classes.

The plenary discussion pointed out that those assigned to observe the pupils should concentrate on them as much as possible and report on them accordingly. They were further reminded to link their reports to the events in the class. It was a useful lesson within a very limited time.

Annex 5. Data Collection Schedule

Region	Teams	Language	District	Date
Northern	A	Dagbani	Nanumba South	4th to 6th June 2017
			Nanumba South	5th to 6th June 2017
		Dagbani	Yendi	7th to 12th June 2017
			Yendi	8th to 12th June 2017
	B	Dagbani	Karaga	5th to 8th June 2017
				5th to 8th June 2017
				5th to 8th June 2017
		Dagbani	Gushegu	9th to 16th June 2017
				9th to 16th June 2017
				9th to 16th June 2017
	C	Dagbani	Tolon	5th to 8th June 2017
				9th to 16th June 2017
	D	Likpakpaln	Yendi	5th to 8th June 2017
				5th to 8th June 2017
				5th to 8th June 2017
		Likpakpaln	Gushegu	9th to 16th June 2017
				9th to 16th June 2017
	E	Gonja	Damongo	5th to 7th June 2017
			North Gonja	7th to 9th June 2017
	F	Brifo	Sawla	4th to 10th June 2017
			Tuna	4th to 10th June 2017

Region	Teams	Language	District	Date
				4th to 10th June 2017
	G	Mampruli	West Mampruli	5th to 14th June 2017
			Mamprugu-Moaduri	5th to 14th June 2017
Upper West Region	H	Dagaare	Wa West	5th to 12th June 2017
			Daffiama-Bussie-Issa	5th to 12th June 2017
			Lawra	5th to 12th June 2017
	I	Sissala	Sissala West	5th to 9th June 2017
Upper East	J	Kusal	Bawku/Binduri	5th to 12th June 2017
	K	Gurune	Bongo/Talensi	5th to 12th June 2017
	L	Kasem	Kasena Nankana W	5th to 12th June 2017
Ashanti		Twi	Afram Plains	5th to 13th June 2017
Brong Ahafo		Twi	Nkoranza	5th to 13th June 2017
		Ewe	Pru	5th to 13th June 2017

Annex 6. Student Sample

IP	Sampling Size	Region	District	Language	Gender: Girls/Boys	No. of centres
Action Aid	210	Northern	Gushegu	Dagbani/Likpakpaln	85/125	14
IBIS	243	Northern	Sawla-Tuna-Kalba	Brifo	108/135	13
ProNet	383	Brong Ahafo	Pru	Ewe	48/49	5
		Northern	West Gonja	Gonja	28/41	3
		Upper West	Daffiamah	Dagaare	22/33	3
		Upper West	Wa West	Dagaa/Brifo	44/77	6
		Upper West	Sissala West	Sissala	23/18	2
CARE International	189	Brong Ahafo	Tain	Twi	21/27	3
		Brong Ahafo	Nkoranza	Twi	19/11	3
		Northern	West Mamprusi	Mampruli	44/67	6
School for life	364	Northern	Yendi	Likpakpaln/Dagbani	104/159	13
		Northern	Nanumba South	Dagbani	63/38	5
GILLBT	76	Northern	Mamprugu-Moaduri	Mampruli	37/39	5
AfriKids	315	Northern	Karaga	Dagbani	50/55	5
		Upper East	Bawku	Kusaal	84/73	7
		Upper East	Bongo	Gruni	11/20	1
		Upper East	Talensi	Gruni	13/9	1
LCD	82	Ashanti	Sekyere Afram Plains	Twi	17/21	2
		Upper East	Kasena-Nankana	Kasem	31/13	2
Plan Ghana	95	Northern	North Gonja	Gonja	21/17	2
		Upper West	Lawra	Dagaare	25/32	3
World Education	45	Northern	Tolon	Dagbani	28/17	3
Total	**2002**				**926/1076**	**107**

Annex 7. Gain Score Regressions Using Implementing Partner in Place of Language

Explanatory variables	Model 1 Basic reading	Model 2 Advanced reading	Model 3 Writing	Model 4 Advanced numeracy	Model 5 Basic numeracy
Age	1.01*	1.95***	1.55**	1.05**	1.86***
Facilitator Assistance	4.94***	2.66	2.14	5.64***	4.97***
Group Work	2.43	4.38**	3.90**	2.93**	3.86**
Pupil language equal LOI	13.23	11.04	8.57	8.98	0.14
Female	-3.39*	-3.79*	-2.38	-1.95	-0.00
Implementing Partner					
Action Aid	-21.26***	-24.86***	-21.71**	-17.16***	-25.27***
AfriKids	-2.62	-9.84	-5.65	-4.60	-8.66
CARE International	-12.60**	-17.35**	-12.56	-8.47	-7.63
GILLBT	-15.72*	-28.51***	-22.10**	-11.01	-15.30
IBIS	4.34	-0.40	0.11	2.28	5.56
LCD	-7.00	-1.45	-6.14	-9.05	-7.80
Plan Ghana	10.91	14.88	7.90	10.90	5.93
ProNet^	0.00	0.00	0.00	0.00	0.00
School for life	-14.04**	-17.85***	-15.26**	-8.62*	-10.27
World Education	-25.62*	-21.91	-22.61	-24.32*	-27.92*
Number of people in household	0.14	-0.01	-0.09	0.09	0.03
Does any activity at home that involves reading or writing	-0.55	0.75	1.77	-0.44	1.52
Has ever attended school?	0.88	2.29	2.13	0.87	-0.71
Number of days missed in the last 5 school days	0.56	0.23	-0.41	-0.24	-0.45
Does any work outside of home	-0.04	-1.73	-1.59	-0.34	-1.04
Wealth Index					
Low^	0.00	0.00	0.00	0.00	0.00
Mid-low	-5.62*	-6.43*	-5.22	-2.31	-1.88
Mid-high	-6.91**	-4.84	-4.37	-3.75	-3.28
High	-3.65	-1.91	-1.44	-2.96	-2.98
Baseline Score	-0.43***	-0.53***	-0.42***	-0.56***	-0.58***
Constant	10.45	5.89	7.58	14.70	14.13
N	1877	1877	1875	1860	1860

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$. ^ = reference group.

All IPs are compared to the reference group (ProNet) which was selected since it is the largest IP.

Bibliography

August, D., & Shanahan, T. (2006). *Developing literacy in second-language students.* Prepared by the Center for Applied Linguistics and SRI International for the Institute of Education Sciences and the Office of English Language Acquisition, US Department of Education; and the US National Institute of Child Health and Human Development. Washington, DC: Lawrence Erlbaum Associates and the Center for Applied Linguistics.

Baroody, A.J., Lai, M.-L. & Mix, K.S. (2006). The development of number and operation sense in early childhood. In O. Saracho, & B. Spodek (Eds.), *Handbook of Research on the Education of Young Children* (pp. 187–221). Mahwah, NJ: Erlbaum.

Brombacher, A. A., Harden, K. K., Nordstrum, L. E., Batchelder, K. E., Cummiskey, C., Dubeck, M. M., & Mulcahy-Dunn, A. D. (2016, July). *Final Findings Report, Tanzania National EGRA*. Prepared for the United States Agency for International Development (USAID).

Chard, J., Clarke, B., Baker, S., Otterstedt, J., Braun, D. & Katz, R. (2005). Using measures of number sense to screen for difficulties in mathematics: Preliminary findings. *Assessment for Effective Intervention 30*(2), 3–14.

Clements, D., & Samara, J. (2007). Early childhood mathematics learning. In F.K. Lester, Jr. (Ed.), *Second Handbook on Mathematics Teaching and Learning* (pp. 461–555). Charlotte, NC: Information Age.

Crown Agents. (2014). *Complementary Basic Education Programme – Year One Learner Assessment Report.* Prepared for the Department for International Development (DFID) and the Government of Ghana.

Directorate of Research, Innovation and Consultancy at the University of Cape Coast (DRIC/UCC). (2016). *Complementary Basic Education (CBE) Learners Assessment: Baseline Report for 2015/2016*. Prepared for the UK Department for International Development (DFID).

Dubeck, M. M., Gove, A., & Alexander, K. (2016). School based assessments: What and how to assess reading. In UNESCO Institute for Statistics (Ed.), *Understanding what works in oral reading assessments* (pp. 41–57). Montreal, Canada: UNESCO Institute for Statistics.

Filmer, D., & Pritchett, L. H. (2001). Estimating wealth effects without expenditure data—or tears: an application to educational enrollments in states of India. *Demography*, *38*(1), 115–132.

Foegen, A., Jiban, C. & Deno, S. (2007). Progress monitoring measures in mathematics: A review of literature. *The Journal of Special Education 41*(2), 121–139.

Gove, A., & Wetterberg, A. (2011). *The Early Grade Reading Assessment: Applications and interventions to improve basic literacy*. RTI International. PO Box 12194, Research Triangle Park, NC 27709-2194.

Graham, S., & Hebert, M. (2010). *Writing to read: Evidence for how writing can improve reading.* A Carnegie Corporation Time to Act Report. Washington, DC: Alliance for Excellent Education. Retrieved from https://www.carnegie.org/media/filer_public/9d/e2/9de20604-a055-42da-bc00-77da949b29d7/ccny_report_2010_writing.pdf

National Center for Family Literacy (NCFL) [US]. (2008). *Developing early literacy: Report of the national early literacy panel. A scientific synthesis of early literacy development and implications for intervention.* Prepared under inter-agency agreement IAD-01-1701 and IAD-02-1790 between the Department of Health and Human Services and the National Institute for Literacy. Washington, DC: National Institute for Literacy. https://www.nichd.nih.gov/publications/Pages/pubs_details.aspx?pubs_id=5750

National Institute of Child Health and Human Development (NICHD) [US]. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups.* U.S. National Institutes of Health Publication No. 00-4754. Washington, DC: NICHD. https://www.nichd.nih.gov/publications/pubs/nrp/Pages/smallbook.aspx

RTI International. (2016). *Early Grade Reading Assessment (EGRA) toolkit, Second edition.* Washington, DC: United States Agency for International Development.

Seymour, P. H. K., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology, 94,* 143–174.

Snow, C. E., Burns, M. S., & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in young children.* Prepared on behalf of the Committee on the Prevention of Reading Difficulties in Young Children under Grant No. H023S50001 of the National Academy of Sciences and the U.S. Department of Education. Washington, DC: National Academy Press.

Snow, C., & the RAND Reading Study Group. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. Research prepared for the Office of Educational Research and Improvement (OERI), U.S. Department of Education. Santa Monica, CA: RAND Corporation.