

WG/GAML/11/2.5.a

BENCHMARK SETTING FOR PRECURSOR SKILLS IN READING: A CASE STUDY OF KENYA





Contents

List of Abbreviations	3
0. Introduction	5
1. Country Profile.....	7
2. Drive towards foundational learning in Kenya: A historical perspective	8
3. Measurements of foundational learning: Results and impact	11
4. Setting National Benchmarks	20
5. Lessons learned from the Kenyan context	30
REFERENCES	34



List of Abbreviations

ADEA	Association for the Development of Education in Africa
AHERI	African Higher Education Research Institute
APBET	Alternative Provision of Basic Education and Training
CBA	Competency Based Assessment
CBAF	Competency Based Assessment Framework
CBC	Competency-Based Curriculum
clpm	Correct letters per minute
CSO	Curriculum Support Officer
Cwpm	Correct words per minute
DfID	Department for International Development
EGRA	Early Grade Reading Assessment
EDC	Education Development Centre
IEAG-SDGs	Interagency and Expert Group on SDGs
IPA	Innovations for Poverty Action
KEYA	Kenya Early Years Assessment
KICD	Kenya Institute of Curriculum Development
KNEC	Kenya National Examination Council
KPEEL	Kenya Primary Equity in Learning Program
Lol	Language of Instruction (LoI)
MSI	Management Systems International
MoE	Ministry of Education

MLA	Monitoring Learner Achievement
MLP	Monitoring Learner Progress
NAC	National Assessment Centre (NAC)
NASMLA	National Assessment System for Monitoring Learner Achievement
NGOs	Non-governmental Organisations
ORF	Oral Reading Fluency (ORF)
PISA	Program for International Student Assessment
PRIMR	Primary Math and Reading Initiative
RTI	Research Triangle Institute
SDG	Sustainable Development Goal
SEE	Sign Exact English
SEACMEQ	Southern and Eastern Africa Consortium for Educational Quality
SNE	Special Needs Education
TAG	Technical Advisory Group
BECF	Basic Education Curriculum Framework
UIS	UNESCO Institute for Statistics



The Report


0. Introduction¹

Despite the concerted efforts by the UNESCO Institute for Statistics (UIS) to enhance reporting on SDG 4.1.1a, an indicator that tracks the percentage of children in grades 2 or 3 achieving minimum proficiency in reading and mathematics, the number of countries reporting on this indicator has remained significantly lower than that of those reporting on primary school enrolment. As a result, the Interagency and Expert Group on SDGs (IEAG-SDGs) downgraded the indicator from Tier I to Tier II, and this decision has raised concerns among stakeholders because it might make countries de-emphasize the importance of addressing learning outcomes in lower primary. To counter this effect, the UIS established a Technical Advisory Group (TAG) to discuss skills that are precursors or correlates for reading comprehension and to determine benchmarks using statistical techniques. UIS is carrying out a global study across many languages and countries to set benchmarks that would be useful to countries. But UIS has also observed that a few countries have already set their own benchmarks using their own methods, and this was thought an interesting alternative to globally-set benchmarks. This Kenya country case study is part of the UIS's engagement to gather information on how individual countries have set benchmarks.

The focus of this case study, therefore, is to describe the process Kenya went through between 2012 and 2019 to come up with lower primary reading benchmarks for both the regular and the Special Needs Education (SNE) learners. As a background to the benchmark setting, the study provides a brief country context, outlining the history of the drive towards foundational learning and discussing existing track record in the assessment of foundational learning at national level. These provide a window into the engagement of the Ministry of Education in Kenya and its development partners with the improvement of foundational learning.

The information reported in this paper was obtained from both document review and interviews with key stakeholders from the Ministry of Education and Non-governmental Organisations (NGOs) supporting the government in the improvement of foundational learning. The key documents reviewed included project reports of the Primary Math and Reading (PRIMR) Initiative, evaluation reports of the *Tusome* intervention, reports of National Assessment System for Monitoring Learner Achievement (NASMLA Class 3), reports of Monitoring Learner Progress (MLP Grade 3), and reports of the benchmark setting process. The preparation too involved interviews with officers from various departments of the Ministry of education, with development partners, with representatives of NGOs focusing on education, and with a professor of Language and

¹ This report was commissioned by the UNESCO Institute for Statistics (UIS) and authored by consultant Prof. Angelina Nduku Kioko.



Literacy Education in the University of Nairobi, whose academic qualifications and research experience are in the area of reading, and who is currently a senior education consultant at the world Bank, a resource person for Association for the Development of Education in Africa (ADEA) and a member of the advisory committee for Innovations for Poverty Action (IPA) and African Higher Education Research Institute (AHERI). From the Ministry of Education, there were three officers from the National Assessment Centre (NAC) in the Kenya National Examination Council (KNEC), two officers from the Kenya Institute of Curriculum Development (KICD), and one officer from the Directorate of Quality Assurance and Standards. From NGOs and development partners, there was one officer from USAID, a senior director from Bill and Melinda Gates Foundation, and one officer from Education Development Centre (which is currently implementing the Kenya Primary Literacy Project focusing on reading remediation). Finally, a University of Nairobi Professor who has been involved in early grade reading research and consultancy for over 20 years also provided input towards this case study.



1. Country Profile

Kenya has made great strides in improving access to education, both at primary and secondary level. In 2003, the government introduced free primary education, and in 2008 free day secondary education. This significantly increased enrolment, and therefore there has been increased need to monitor the achievement of learning outcomes as resources get overstretched by the growing enrolment.

At the time of this case study, Kenya was close to the tail end in implementing basic education reforms aimed at aligning the education system to: the Constitution of Kenya 2010, Kenya Vision 2030, the East African Community Curriculum Harmonization Structures and Framework, and other policy documents that express the aspirations of the country. These reforms provide a comprehensive conceptualisation of basic education as: pre-primary (2 years), primary 6 years), lower secondary – Junior School (3 years) and Senior School (3 years). Inclusive education forms a core part of these reforms; thus, the Basic Education Curriculum Framework (BECF) provides for two pathways for learners with special needs – the age-based pathway and stage-based pathway. Learners with special needs and disability follow either pathway depending on the nature and severity of the disability. These education reforms are simply referred to as the Competency-Based Curriculum (CBC) and they adopt a learner-centred and skills-oriented approach, introducing flexible pathways and emphasising early identification of talents. In addition to summative assessment at the end of each level, the system also incorporates continuous assessment to track learners' progress through school-based assessments, whose results are uploaded to the National Examination Council portal.

With regard to reading outcomes, Kenya is a multilingual country with two official languages (Kiswahili and English) and close to 70 indigenous community languages. Though for a long time, English has remained the Language of Instruction (LoI) in basic, tertiary and university education, the other languages have a role to play in the education system. Kiswahili, the second official language and also the national language, is taught as a compulsory language throughout the basic education levels, and, because in the rural areas most children join school without either Kiswahili or English, the policy/practice has been to use “the language of the catchment area” as LoI up to the end of lower primary (Grade 3). In the CBC reforms, the lower primary school timetable also has two lessons per week assigned to the teaching/learning of an indigenous community languages, and these indigenous languages remain optional learning areas in upper primary school. Therefore, at the end of lower primary school, a majority of Kenyan learners are exposed to literacy instruction in at least three languages at various levels, which can pose a challenge on the speed at which literacy skills in English develop.




2. Drive towards foundational learning in Kenya: A historical perspective

The education system of a country needs to provide evidence that learners have mastered the basic literacy, numeracy, and critical thinking skills essential for further learning and daily functioning, and that these basic skills are acquired at an early stage in the learning process. For the last over 20 years, the engagement with improvement of foundational learning in Kenya has been driven by the research findings that indicated that children were not learning. As a member of the global community that made a collective commitment in the Dakar Framework for Action in April 2000 to achieve six goals by 2015, and particularly in response to goal number six of this framework, on “improving all aspects of quality education and ensuring excellence of all so that recognized learning outcomes are achieved by all especially in literacy, numeracy and essential life skills”, actors in Kenya started engaging in research to provide the country with evidence towards this end. Much of the early evidence was shocking. For example, from 2009, Uwezo Kenya carried out household surveys to monitor learning outcomes on literacy and numeracy among children aged 7-13 years, regardless of grade. These surveys revealed very low achievements. For example, the *Uwezo Annual Assessment Report* (2011) study indicated that only 28% of the children in Class 3 (the end of lower primary) could read a Class 2 text, and 20% of learners in Class 5, and 4% of learners in class 8, (the last class in primary school at that time) could not read a class 2 level text (Mugo, et al., 2011).

2.1 The Primary Math and Reading (PRIMR) Initiative

At the same time, a USAID-funded research program, the Primary Math and Reading (PRIMR) Initiative undertook a baseline as a basis to evaluate the intervention. The focus of the initiative was to answer the questions: “Can Kenyan children read? Can they comprehend? Is Kiswahili or English easier to learn to read? Do Kenyan children have basic Maths skills? Can reading and Maths skills be improved?” (USAID, 2012, p.1). The baseline findings were based on a survey of learners in Class 1 and 2 in 220 schools randomly selected from the 500 schools in which the project would run. A total of 4,385 learners were assessed: 2,192 class 1 learners and 2,193 class 2 learners, in a gender-balanced sample- 2,186 girls and 2,199 boys. For reading, the assessments included a variety of EGRA subtasks, letter sound fluency, non-word fluency, oral reading fluency, reading comprehension, and listening comprehension. The findings indicated that on the average, across Class 2, only 5.1% and 7.2% of the learners could fluently read Kiswahili and English respectively. Fluency was defined on the country reading benchmark of 45 cwpm in Kiswahili and 65 cwpm in English, (USAID, 2012, p.2). This confirmed the findings by the Uwezo surveys and indicated the need for urgent intervention at this level.

The one-year review of the PRIMR initiative indicated that the project had positively impacted the oral reading fluency benchmark achievement, and reading comprehension, though effects varied by language, grade, and school type. Statistically significant improvements in oral reading fluency were observed in grade 1 and grade 2 non-formal schools and grade 1 formal schools for both English and Kiswahili, with the greatest impact being in grade 1 non-formal schools and in English. Overall, more learners were reading at




the national fluency benchmarks than at baseline (Piper & Mugenda, 2012). The PRIMR endline evaluation report in 2014 indicated that PRIMR significantly improved letter-sound fluency, with treatment pupils scoring 47.0 clpm compared to 25.7 correct letters per minute (clpm) for control pupils, a causal effect of 21.3 clpm (0.73 SD). In oral reading fluency, the effect was 13.7 cwpm overall, representing over one year of learning gains relative to control schools. Reading comprehension scores in Class 1 were 21.1% in PRIMR compared to 9.8% in control schools, while the Class 2 gain was 17.3%. By the endline, over twice as many learners in the PRIMR treatment in Classes 1 and 2 met the KNEC English benchmark (65+ cwpm), with moderate to large effect sizes averaging 0.46 SD (USAID, 2014).

2.2 The *Tusome* intervention

As the PRIMR initiative came to an end in 2014, the Kenyan government decided to scale up the intervention to all public schools in Kenya and named the intervention *Tusome* ('let us read' in Kiswahili). The decision to scale up the gains of the PRIMR project was cited by nearly all the officers interviewed for this study as the most comprehensive drive by the Kenyan government to improve foundational literacy. In effect, the Ministry of Education was scaling up the following defining features of the PRIMR initiative:

- a) Evidence-based structured approach for teaching reading in English and Kiswahili.
- b) Extensive training and support of teachers through coaching and regular classroom observations to ground effective reading instruction techniques, including the use of decodable texts and comprehension strategies.
- c) High-quality learning materials which included levelled readers and textbooks with graduated decodable texts to ensure that learners had access to engaging, developmentally appropriate reading materials in English and Kiswahili.
- d) Use of technology to collect real-time data for monitoring learner progress and teacher performance, allowing for timely feedback and adjustments to the program.
- e) Assessment based on the Ministry of Education established reading fluency and comprehension benchmarks to track learner progress and to ensure accountability and measurable impact.
- f) Engagement of already established/existing Ministry of Education structures.

After an intense preparation period which involved the training of trainers at the national, county, sub-county, zonal, and school levels; and the revision and provision of the PRIMR learners' books and teachers' guides, the *Tusome* Early Grade Reading intervention in Kenya was implemented in January 2015. It was implemented nationwide, involving over 20,000 public schools and 8,500 private schools, and was integrated into the Ministry of Education systems for long-term sustainability. It was funded by USAID/Kenya and DFID/Kenya and executed by the Kenyan Ministry of Education with technical support from RTI *Tusome*. The provision of the *Tusome* learners' books targeting a student book ratio of 1:1 and teachers' guides with scripted lessons to guide the structured direct method of reading instruction, and the robust in-service teacher training, teacher coaching, and lesson observation/monitoring reached a level of engagement that had not been witnessed in the country before.



Though the implementation of educational reforms before the completion of the initial three years of the supported *Tusome* intervention, and the closure of schools for a long period during the COVID-19 pandemic negatively affected the success of this intervention, there are several gains from this intervention that will continue to impact the development of foundational literacy in Kenya positively for a long time. The *Tusome* learners' books and teachers' guides have gone through significant changes as they were revised to align to the CBC curriculum, but the structured instruction and the effort to organise reading instruction systematically from simple decodable texts to more complex texts have been retained in the learning materials to a larger extent. Further, the fact that this approach has been built into the curriculum designs for lower primary school ensures a good level of sustainability of the change in classroom instruction.

The *Tusome* programme also has had lasting impact on the way the Ministry monitors learning assessment and provides teacher support. Under *Tusome*, an education dashboard was developed to collect student learning and teacher performance data that was managed and maintained by the Ministry of Education. Through a partnership with Google, *Tusome* equipped Curriculum Support Officer (CSOs) and Quality Assurance Support Officer with tablets loaded with open-source Tangerine software and trained them to assess learners' reading abilities using forms in the tablets. These officers would regularly visit schools and upload learner assessment data to the education dashboard, enabling the Ministry of Education at county and national levels to track learner progress nationwide in real time. The officers would also use the tablets to conduct regular classroom observations, provide instructional feedback to teachers, and upload observations and feedback to the education dashboard, strengthening their accountability for teacher performance.

Because the *Tusome* baseline, midline and endline studies provide data on the status of foundational literacy nationwide, the assessment arm of the programme is further discussed in the next section on 'measurements of foundational learning'.

2.3 Way forward on the drive towards foundational learning

Implementing the *Tusome* intervention nationwide was seen as a great move and, as will be shown below, the achievements at midline were significant. For Kenya to recover the gains lost as a result of: a) implementation laxity after the project ended, b) the impact of the implementation of curriculum reforms in the middle of the intervention, and c) the effect of school closure during the COVID-19 pandemic, there is need to establish systems of re-engineering teacher support, supporting catch-up programmes for learners who were caught in the middle of these changes, and reviewing the English and Kiswahili pre-service teacher training designs need revision to prepare teachers on early literacy pedagogy.



3. Measurements of foundational learning: Results and impact

The government of Kenya underscores the importance of conducting assessment of the education system at various levels nationally, regionally and internationally to monitor learners' achievement of the knowledge, skills and values spelt out in the national basic education curriculum. Kenya institutionalised the National Assessment System for Monitoring Learner Achievement (NASMLA) in 2006 to undertake Monitoring Learner Achievement (MLA) studies at the Basic Education level and created the National Assessment Centre (NAC) in the Kenya National Examination Council (KNEC) with a mandate to provide leadership in these large-scale assessments. The revised Monitoring Learner Achievement Framework for Basic Education (available on the NAC website) proposes to undertake five assessments for the age based pathway at national level: a) School readiness survey at the end of pre-primary (5-year-olds); b) NASMLA Grade 3 at the end of lower primary (8-year-olds); c) NASMLA Grade 6 at the end of primary (11-year-olds); d) NASMLA Grade 8 in the middle of junior school (13-year-olds); and e) NASMLA Grade 11 at the start/middle of senior school (15-year-olds). In addition, the framework provides for one regional assessment – Southern and Eastern Africa Consortium for Educational Quality (SEACMEQ) at Grade 6 and one international assessment – Program for International Student Assessment (PISA) for 15-year-olds. The NASMLA assessments aim at gathering empirical data on the level of mastery of specific subject skills and competencies at formative stages as learners progress in the school system and in order to inform the country on the quality of education by reporting on factors such as pupil, teacher, school and home characteristics that influence the achievement of learning outcomes.

At national level, there are two main sources of assessment data on lower primary schooling: the direct Kenya National Examination Council assessment programs, and those conducted by the Ministry of Education development partners. In this section we will review assessment data from two KNEC led assessments, NASMLA Class 3, KEYA/MLP Grade 3; and four sets of assessment data from the *Tusome* intervention.

3.1 National Assessment of learner achievement by KNEC

At lower primary KNEC has two assessments targeting end of lower primary: the NASMLA Class 3 which has had three reports posted on the website - 2010, 2016, and 2018; and the Kenya Early Years Assessment (KEYA) which has two reports on the website: MLP 2019 and MLP 2022. The staff interviewed indicated that there is an on-going assessment supported by the on-going Kenya Primary Education Equity in Learning project (KPEEL), which adopted Grade 3 MLP - 2022, as the baseline, collected midline data in 2024 (report to be available in 2025), and plans an endline study in 2026.

b) NASMLA Class 3

The reports for the NASMLA 2010, 2016 and 2018 are available on the NAC website (<https://nac.knec.ac.ke>). These are national sample studies, and the sampling is conducted

using a sampling frame that ensures national representativeness (see the methodology section in KNEC, 2010, 2016 & 2020). The NASMLA studies in 2010, 2016 and 2018 set the benchmark for minimum proficiency as 50% score in the items in the assessment tool and adapted the four competency levels used in the SEACMEQ studies to categorise learner performance: pre-reading, emergent reading, basic reading, and reading for meaning. For each of the levels, a descriptor of the competency is provided, and the test items aligned to each of these competency levels are identified. Level 1 is conceptualised as the most basic and level 4 as the most advanced as shown by the descriptors in the table below.

Table 1: Descriptors for the 4 competency levels in NASMLA studies


LEVEL	DESCRIPTOR
1- Pre-reading	Matches words and pictures involving concrete concepts and everyday objects.
2- Emergent reading	Spells correctly simple everyday words and recognizes missing letters in such words. Uses familiar words to complete simple everyday sentences.
3- Basic reading	Uses correct punctuation in simple sentences. Infers meaning from short passages and interprets meaning by matching words and phrases. Identifies the main theme of a picture.
4- Reading for meaning	Links and interprets information located in various part of a short passage. Understands and interprets meaning of a picture and writes short sentences to describe the theme.

Generally, these studies revealed low achievement of basic literacy skills among Class 3 pupils as indicated on Table 2 below (Note that NASMLA Class 3, 2010 assessed literacy in English only.)

Table 2: NASMLA Class 3: Proportion of learners achieving various competency levels

READING LEVEL	ENGLISH			KISWAHILI		
	2010	2016	2018	2010	2016	2018
1- Pre-reading	6.2%	14.9%	13.1%	-	10.5%	12.5%
2- Emergent reading	46.1%	60.1%	63.3%	-	63.1%	77.7%
3- Basic reading	36.7%	38.1%	53.1%	-	69.2%	70.4%
4- Reading for meaning	11%	28.6%	41.2%	-	47.1%	48.7%

(KNEC, 2010, p.80; 2016, p.115; 2020, p.194)



Interpreting the figures and the reporting narrative in the NASMLA Grade 3 reports proved very challenging, which confirms the need for technical capacity building in data analysis and research reporting as was highlighted in the interviews with officers from NAC.


Though the figures for 2010 add to 100%, those for 2016, and 2018 (reported in 2020) do not, thus the meaning of the percentages is not clear from what is presented in the documents. Interviews with officers from NAC indicated that, for example, in the 2010 report, 6.2% of the learners could only achieve at level 1, while only 11.0% had level 4 competence. From the report, the “desired” were levels 3 and 4, which meant that 53.3% of the learners were not meeting expectations (KNEC, 2010, p.80). The interview further indicated that in the 2016 and 2018 figures, the proportion of learners at each level is calculated out of a hundred. Thus, in English only 14.9% and 13.1% of the learners demonstrated achievement at level 1 (pre-reading) in 2016 and 2018 respectively. This is the reverse of the expected! In the interview, the NAC officers indicated that there were misplaced items in both the English and Kiswahili tools which explain the low achievement at level 1 in both languages. With that reading, the figures in Table 2 show steady improvement in the proportion of learners performing at levels 3 and 4.

Both NASSMLA 2016 and 2018 (reported in 2020) studies revealed gender disparities in the achievement of literacy skills among Class 3 learners. For example, in the 2018 study, the proportion of girls attaining the 50% score benchmark in Kiswahili and English was higher than that of boys at 71.3% of the girls against 64.3% of the boys, and 58.5% of the girls against 47.4% of the boys respectively (KNEC, 2020).

Responses to the interview questions did not provide specific examples on how the findings of the NASMLA Class 3 assessments have influenced education policy and practice in the country. Though all the officers who were interviewed seemed convinced that these findings have informed interventions in early grade literacy, no specific reference to policy or intervention was directly linked to NASMLA Class 3 assessments. One of the respondents indicated that the 2010 NASMLA study served as a baseline for many NGO interventions and may have been one of the motivators for the Primary Math and Reading (PRIMR) Initiative.

c) Monitoring Learner Progress (MLP) Grade 3

The other significant measurement of foundational learning in Kenya, conducted directly by the Kenya National Examination Council is the Kenya Early Years Assessment (KEYA) at Grade three and currently reported as Monitoring Learner Progress (MLP) Grade 3. This is a census assessment administered annually to all learners in all schools that have enrolled Grade 3 learners in Kenya. With the implementation of CBC, the Ministry of Education through the Kenyan National Examinations Council (KNEC) developed the Competency Based Assessment Framework (CBAF) to guide the assessment of the new curriculum. The framework distinguishes assessment for learning (formative) and assessment of learning (summative) to reflect the paradigm shift in pedagogy, through the use of the Competency Based Assessment (CBA) approach. CBA focuses on actual skills and knowledge, is criterion-referenced and aligned to the learning outcomes outlined in the curriculum designs.




Assessment is administered at different levels of the basic education using a variety of assessment tools to cater for the diversity of learners.

At the end of lower primary (Grade 3), KNEC administers the Kenya Early Years Assessment (KEYA) - a national assessment which assesses three learning areas: Literacy (English), Numeracy (Mathematics) and Integrated Learning Areas for the learners pursuing the age-based pathway. KEYA mainly focuses on establishing levels of competencies in the learning outcomes and is not used to determine transition to Grade 4. The assessment tools and scoring guides are developed by KNEC through a rigorous process and are uploaded on the KNEC Assessment Portal for schools to download and administer. The administration and scoring are done by the classroom teachers at the school level, then the schools upload the scores onto the KNEC CBA portal for verification, validation, analysis, and interpretation.

Once KNEC finalises the verification, validation, analysis and interpretation, the report is disseminated to stakeholders and posted on the KNEC website as Monitoring Learner Progress (MLP) Grade 3. At the time of this case study, two of these reports (KNEC, 2019 & 2022) were available on the website (<https://nac.knec.ac.ke>). The results for literacy (English) are presented using categories based on the strands in the curriculum designs. The 2019 MLP Grade 3 assessed learner achievement using items that were grouped into four tasks: listening and speaking /signing skills; reading aloud/ signing; reading comprehension skills and language structures; and writing skills. Learner performance in each of these tasks was scored using four levels: Below Expectation (level 1), Approaching Expectation (Level 2), Meeting expectation (level 3) and Exceeding Expectation (level 4). The proportion of learners performing at level 3 and above in all the tasks was commendable: listening and speaking were at 65%, reading aloud at 61%, reading comprehension and language structures at 51%, and writing at 47%. This, though, meant that 35%, 39%, 49%, and 53% of learners performed below 'meeting expectation' (Level 3) in listening and speaking, reading aloud, reading comprehension and language structures, and writing respectively. This would be the proportion of learners in need of intervention.

The 2022 MLP reported on the same skills but grouped the tasks differently to separate reading comprehension and language structures. The proportion of learners performing at level 3 and above nearly matches the 2019 performance in all the tasks: listening and speaking were at 65.3%, reading aloud at 60.3%, reading comprehension at 55%, grammar at 54%, and writing at 44%. This flags out 35%, 40%, 45%, 46% and 56% of learners as below 'meeting expectation' (Level 3) in listening and speaking, reading aloud, reading comprehension, grammar, and writing respectively. Though the 2024 report was not accessible at the time of this case study, the scoring guide for that year was accessed from the primary schools, and it provided information on the ORF benchmark used in KEYA as 90 cwpm – the level 3 of performance.




The MLP reports indicate that the outcome of the assessment is utilised for improvement at the various levels in the education system. At the school level, it provides feedback to learners on their levels of achieving the competencies in foundational learning and at the same time informs teachers who will receive the learners in Grade 4 on the exit behaviour of the Grade 3 learners. At the national level, the MLP is used to inform policies regarding CBC implementation, specifically in curriculum development, teacher capacity and other policies related to quality such as: “monitoring quality of education in the early years of education; gathering evidence on learning achievement at the end of Early Years Education; establishing the interventions required for teacher capacity and support; enhancing school and educator accountability; improving learning and instruction; informing education planning and policy; and informing quality of curriculum, including curriculum load, appropriateness, difficulty” (KNEC, 2019, p.1).

Even though both NASMLA and MLP are KNEC administered assessments, currently the two are not harmonised in terms of how they categorise the skills of literacy they are measuring. While NASMLA works with 4 levels of reading competency and reports on the proportion of learners successfully reading at each of these levels, MLP follows a language skills pattern and reports on the proportion of learners at the various competency levels in listening, speaking, reading, grammar and writing. KNEC does not make an explicit effort to relate NASMLA and MLP (perhaps because they belong to two different education systems), and though the ORF in MLP is reported using four performance levels, there is no clear indications that learners performing at Level 1 (reading less than 45 cwpm) can be said to comparable abilities with learners at Level 1 (‘Pre-reading’) in the NASMLA studies. The two measurements appear to run independent of each other. Further, the reports do not end with specific ‘actionable’ recommendations for improvement which would ensure meaningful expectations or target setting in the next assessment. For example, there no clear targets of success set between one KEYA and the next to help evaluate the impact of interventions in between.

3.2 Learning assessment by the *Tusome* intervention

As indicated earlier, the *Tusome* baseline, midline and endline studies provided data at national level on the proportion of early grade learners meeting the national reading benchmarks at various levels in both English and Kiswahili. Though the number of research questions at the baseline were fewer than those at midline and endline evaluations, two key questions which speak directly to early grade literacy remained constant and guided the measurement of reading achievement: 1) What are the levels of Classes 1 and 2 learners on the reading subtasks? 2) What proportions of Classes 1 and 2 learners can read grade-level text? An external evaluation of the programme was conducted by Management Systems International (MSI). The baseline (pre-test), midline (first post-test) and endline (second post-test) assessments were sample-based assessment of the reading skills of Class 1 and 2 learners. Using a three-stage cluster sampling procedure from a sampling




frame, 26 counties were sampled out of the total 47; 204 schools comprising 174 public schools and 30 Alternative Provision of Basic Education and Training (APBET) out of the total 22,154 public schools and 1,000 APBET schools; and 24 pupils per school, with 12 (6 boys and 6 girls) in each of the Classes: 1 and 2. This gave a target of 4,896 total learners - 2,448 boys and 2,448 girls for the assessment (USAID, 2016, p.3). The report indicated that the actual numbers of schools and pupils were close to the target numbers. All 204 schools were reached during data collection. A total of 4,866 pupils were tested (99 percent of the target). Prior to the data analysis, the MSI statistician applied sampling weights to the data so that the data sets would be nationally representative².

The *Tusome* assessments used EGRA tools, adapted to measure reading skills in English and Kiswahili in the Kenyan context. The English version had eight tracks: phoneme segmentation (untimed), letter sound knowledge (timed), invented/non-word decoding (timed), Vocabulary (untimed), Passage reading (A-timed), reading comprehension (A-untimed), passage reading (B-timed), reading comprehension (B); while the Kiswahili one had six tracks: letter sound knowledge (timed), invented/non-word decoding (timed), Passage reading (A-timed), reading comprehension (A-untimed), listening comprehension (untimed), syllable fluency (timed).

In addition to reporting on the proportion of learners performing at each of four reading categories of the English and Kiswahili fluency benchmarks: zero-readers, beginning readers, emergent readers, and fluent readers, the reports also present learner performance per the sub-tasks. At baseline in 2015/2016 for Class 1 English, 50.3% learners were at zero reader level, 35.6% at beginning, 11.2% at emergent, and only 2.9% at fluent reader level while in Class 2, 35.8% were at zero- reader level, 27.6% at beginning, 23.1% at emergent and 13.3% at fluent level. In Kiswahili, the baseline assessment reported 67.6%, 18.2%, 13.4%, 0.8% zero readers, beginning, emergent and fluent readers respectively in Class 1; and 41.3%, 18.9%, 35.0%, and 4.8% zero, beginning, emergent and fluent readers respectively in Class 2 (USAID, 2015/16).

The midline study was conducted in 2016 and reported in in 2017. At midline, 30% of Class 1 pupils were emergent readers, and 18% fluent readers, while 29% of Class 2 pupils were emergent readers and 47% fluent readers in English (USAID, 2017). Thus, English reading performance improved between baseline and midline, with a lower percentage of pupils in the zero and beginning reader categories and a higher percentage in the emergent and fluent reader categories. The same was noted for Kiswahili at midline: 32% of Class 1 pupils were emergent readers and 3% percent fluent readers, while 54% of Class 2 pupils were emergent readers and 12% fluent readers. Thus, in Kiswahili, the proportion of learners at zero and beginning reader categories reduced significantly while the proportion of emergent and fluent readers increased between baseline and midline (USAID, 2017).

² The training of enumerators, the data collection and analysis process are fully described in the reports cited here.



However, the large drop in non-readers (or zero readers) observed between baseline and midline was partially reversed by endline in 2019 (reported in 2020) for both grades and in both languages, as was the increase in fluent and emergent readers. The findings of the endline for Class 1 English were: 34% at zero reader level, 38% at beginning, 14% at emergent and only 14% at fluent reader level while in Class 2, 14% were at zero-reader level, 32% at beginning, 35% at emergent and 18% at fluent level (USAID, 2020). In Kiswahili, endline reported 62%, 11%, 22%, 4% zero readers, beginning, emergent, and fluent readers respectively in Class 1; and 29%, 17%, 46%, and 8% zero, beginning, emergent and fluent readers respectively in Class 2. Thus, only 18% and 8% of grade 2 learners met the country's minimum grade-level reading proficiency benchmark in English and Kiswahili respectively, by endline; and the proportions of zero-readers in both languages were higher at endline than had been at midline (USAID, 2020).

The interviews identified a number of factors that may have led to this reverse of progress: 1) a number of the teachers who had been trained moved to upper primary with their learners and those who took over the lower primary had not undergone training for *Tusome* pedagogy; 2) the teacher coaching/support weakened with changing responsibilities of CSOs; 3) teacher training for the implementation of CBC departed from some of the *Tusome* classroom routines; 4) the implementation of the CBC reforms which reduced the number of lessons for both English and Kiswahili.

In June 2021, the Kenya National Examinations Council (KNEC), in collaboration with *Tusome*, undertook an Early Grade Reading Assessment (EGRA) study to establish learning outcomes for grades 1, 2, and 3 with an aim of comparing these literacy outcomes in 2021 with those recorded during the *Tusome* external evaluation done in 2019 (USAID, 2022). They hoped to make recommendations to education stakeholders on effective strategies to halt the learning loss experienced during the COVID-19 school closures and support the MoE in designing interventions to ensure long-term improvements in literacy outcomes across the country and support the portions of the country that are struggling the most with literacy outcomes.

To measure the learning outcomes, the study administered EGRA instruments in English and Kiswahili to Grades, 1, 2 and 3 learners in 210 schools: 180 public, 30 APBET (a representative national sample). Data was collected using Tangerine software on tablets and analysed using Stata statistical analysis software. The report provides detailed performance per the EGRA sub-tasks and also reports the proportion of learners falling in the various categories with reference to the country reading benchmarks for these grades and compares the findings with those of *Tusome* endline in order to determine whether there has been progress. The comparison of the performance with the *Tusome* endline study is presented in Table 3 below:

Table 3: Proportion of learners reading at the various categories: *Tusome* endline 2019 and the EGRA study 2021

CATEGORY	ENGLISH					KISWAHILI				
	2019 ³		2021			2019		2021		
	G-1	G-2	G-1	G-2	G-3	G-1	G-2	G-1	G-2	G-3
Zero-readers	34%	14%	45%	27%	12%	62%	29%	65%	35%	16%
Beginning readers	38%	32%	35%	30%	23%	11%	17%	12%	21%	29%
Emergent readers	14%	35%	10%	32%	37%	22%	46%	17%	37%	43%
Fluent readers	14%	18%	10%	12%	27%	4%	8%	6%	7%	12%

From Table 3, it is clear that there are more zero-reader in both English and Kiswahili in 2021 than there were at the end of the *Tusome* intervention in 2019. Further, the proportion of emergent and fluent readers in both Grades 1 and 2 went down. For example, combining emergent and fluent readers in English, the proportion went down from 28% in 2019 to 20% in 2021 in Grade 1, and from 53% in 2019 to 44% in 2021 in Grade 2. Thus, the number of Grade 2 learners who could not read a single English word correctly nearly doubled from 14% in 2019 to 27% in 2021, and the fluent readers dropped from 18% in 2019 to 12% in 2021.

One of the clear explanations of the erosion of the *Tusome* national gains was the impact of the pandemic which resulted in a nine-month closure of all schools in 2020 and an abbreviated academic calendar that condensed three academic years (2020, 2021, and 2022) into two calendar years (2021 and 2022), further reducing classroom instruction by as much as 40 percent. Also, as indicated above the implementation of the CBC reforms initially reduced classroom instruction time for both English and Kiswahili.

³ The figures for 2019 are taken from the *Tusome* endline study, and for 2021 from the EGRA Study 2021.



3.3 Way forward in measurement of early grade learning achievements

As the discussion above shows, Kenya is rich in terms of the presence of assessment data of foundational literacy skills. However, these reports do not ‘speak the same language’ and may not help the country in reporting internationally in a systematic way. For example, while one of the KNEC assessments reports using achievement levels adapted from SEACMEQ, the other KEYA makes use of levels aligns to the language skills in the curriculum. The *Tusome* assessments, on the other hand, use adapted EGRA tools and focus on reporting on performance of learners in each of the EGRA sub-tasks, and ultimately report on the proportion of learners reading at the national ORF fluency benchmarks (based on fluency that matches acceptable comprehension) across the four categories of Zero reader, beginning reader, emergent reader, and fluent reader in the various grades assessed. There is, therefore, need for the country to harmonise the system of assessing and reporting learner achievement and ensure that this system is followed whether the assessment is conducted by the ministry of Education development partners or by NAC, the section in KNEC mandated to conduct national assessments.



4. Setting National Benchmarks

The answer to the question why Kenya developed early grade reading benchmarks at the time it was done is provided by the response below from one of the interviewees for this report:


“The system needs an understanding of whether kids are learning. Are they or are they not on track? We needed a system, a simple way to look at it that doesn't depend on data from outside of Kenya and outside of the region. So, the advantage that we had was we had the availability of data that was Kenyan, data that was active, data that was nationally representative to help develop these benchmarks that could help with policy makers' decision making.” Dr. Ben Piper.

The review of documents and interviews with both Ministry of Education official and education development partners revealed the existence of two sets of national reading benchmarks for end of lower primary, which have been developed at different times for different purposes. One set is variably referred to as the national benchmarks, the KNEC benchmarks, the Ministry of Education benchmarks, the *Tusome* benchmarks. These reading benchmarks cover Grades 1-3 in both Kiswahili and English, and blind and deaf learners, and are said to have been developed measure the interventions of the *Tusome* programme. These benchmarks are widely referenced in reports on the Ministry of Education scale up of early grade reading intervention – *Tusome* and in NGO projects focusing on early grade reading. The second set of national benchmarks is invariably referred to as the ‘KNEC benchmark’ and is referenced in the curriculum designs and the latest assessment of Grade 3 in CBA, the Kenya Early Years Assessment (KEYA). There was more information on the process that led to the development of the national benchmarks to measure the *Tusome* intervention, therefore, this case study will first focus on the process used to set those benchmarks and provide a brief on the second set at the end.

4.1 The benchmarks developed during the *Tusome* intervention

The first set of national lower grade reading benchmarks in Kenya in both English and Kiswahili, and for blind and deaf learners are explained as “the level of reading fluency (measured in correct words per minute (cwpm)) at which comprehension is acceptable”. Thus, though they are pegged on fluency, they take into account comprehension. The analysis that was used to align fluency and comprehension considered the linguistic and orthographic differences between Kiswahili and English and set different benchmarks for the two languages. Based on the number of correct words read by a learner in a connected text over a duration of one minute and the point at which this reading converges with acceptable comprehension, benchmarks for Oral Reading Fluency (ORF, measured in cwpm) for foundational literacy in both English and Kiswahili for regular learners, blind learners and deaf learners were arrived at.

The initial process to develop benchmarks was supported by the Primary Math and Reading (PRIMR) Initiative in 2012. The project provided technical support to the Kenya National Examinations Council (KNEC) on the establishment of benchmarks on literacy in English and Kiswahili for grades 1 and 2 for the regular school-going learners. These initial benchmarks



are widely referenced and used in the PRIMR baseline, midline and endline studies. They formed the measurement by which the PRIMR project evaluated learner progress during the implementation of the project and interpreted. However, the documents describing the process by which these benchmarks were arrived at were not accessible during the development of this Kenya case study, thus, the need to rely on key informant and expert interviews as a way to reconstruct the history of these processes. From the interviews, it was indicated that the Kenya National Examination Council led this initial process and Research Triangle Institute (RTI) provided technical support using EGRA data obtained in the baseline PRIMR study.

Towards the end of the PRIMR project, the government of Kenya, with the support of USAID and Department for International Development (DfID) funds, and technical support from RTI decided to scale up the intervention to all lower primary schools in Kenya in what came to be known as *Tusome*. This meant that the intervention would spread to cover Grade 3 and schools for the deaf and blind which had not been covered in the initial benchmarks. Secondly, because the data sample that had been used during the development of the 2012 benchmarks (for Grade 1 and 2) were project specific and not nationally representative, it was necessary to review those benchmarks using nationally representative data. This was seen as one of the ways to ensuring the sustainability of the scale up.

The process involved the engagement of relevant stakeholders through workshops and session briefings. Key engagements included: a) an initial meeting with the KNEC technical committee to discuss and develop a shared understanding on the need to develop the benchmarks, and agree on the process of developing them; b) a one day workshop with stakeholders and KNEC technical committee members to discuss and develop draft benchmarks; c) a briefing meeting with the Chief Executive Officer (CEO) KNEC; and d) a workshop with the KNEC steering committee for presentation, discussion and adoption of the draft benchmarks.

4.1.1 Meeting with KNEC Technical Committee

The RTI team organized a planning meeting with the KNEC Technical Committee team at the KNEC Offices in January 2019. This was a planning meeting during which the two parties planned for the benchmarking workshop. They agreed on the content of the presentation to be given during the benchmarking workshop, on proposed participants to be invited and on the roles to be played by KNEC and RTI-*Tusome* during the benchmarking workshop. The team also needed to define the reading precursor skills and tasks for which benchmarks were to be developed and discussed in the upcoming workshop (letter naming fluency, oral reading fluency, reading comprehension). They agreed that there would be need to consider more than one benchmark level, so that there would be a lower and an upper benchmark for each grade.

4.1.2 Workshop with Education Stakeholders

The workshop with education stakeholders was held in February 2019. The organizing team was the National Assessment Centre (NAC) - a centre within the Kenya National Examination Council in charge of the National System for Monitoring Learner Achievement

(NASMLA), with technical support from RTI *Tusome*. The workshop participants included representatives from various sectors in the Ministry of Education, from RTI *Tusome*, from the USAID, and from one NGOs engaged in education. A total of 46 stakeholders (20 male and 26 female) participated in the workshop. Table 4 provides details on workshop participants.

Table 4: Stakeholder Workshop Participants

Organization	Male	Female	Total
Access Education International (NGO)	1	0	1
Aga Khan Primary School	0	1	1
Kenya Society for Deaf	1	0	1
Kenyatta University	0	1	1
Kenya Institute for the Blind	0	1	1
Kenya Institute of Curriculum Development	1	1	2
Kenya National Association of the Deaf	2	0	2
Kenya National Examination Council (KNEC)	2	5	7
Kenya National Union of Teachers (KNUT)	1	0	1
Kenya Publishers Association	1	0	1
Machakos Teacher Training College	0	1	1
Ministry of Education (MOE)	2	6	8
Park Road Primary school	0	1	1
RTI- <i>Tusome</i>	4	3	7
Teachers Service Commission (TSC)	2	1	3
Twaweza East Africa (NGO)	1	0	1
University of Nairobi	0	1	1
United States Agency for International Development (USAID)	2	2	4

Organization	Male	Female	Total
United States International University	0	1	1
World Vision – Kenya (NGO)	0	1	1
Total	20	26	46

(Development of Benchmarks on Early Grade Reading (Report availed by Dr. Ben Piper)

In the introduction, the coordinator of the National Assessment Centre at KNEC linked the objectives of the workshop to the Sustainable Development Goal number 4, which she explained required countries to set their minimum proficiency levels. She noted that in the absence of set minimum proficiency levels for primary schools in Kenya, assessments have historically assumed that a 50% is score is the benchmark. The coordinator explained that the specific objectives of the workshop were to review the benchmarks for grade 1 and 2 and set the benchmark for grade 3 learners and the Deaf and Blind learners through interrogation of representative data on performance.

The workshop participants were then taken through the process of benchmark development by a senior representative from the RTI *Tusome* team. He explained the role of benchmarks in guiding instruction, monitoring progress, setting performance expectations in Grades 1–3 in both regular and special schools, and anchoring assessment. He pointed out that benchmark setting is systematic and should be data-driven so that the standards established can guide teaching, learning, and assessment practices. He then shared the findings of the *Tusome* midline study (Grade 1 and 2), and APBET baseline (Grade 3) where the dataset reliability scores for ORF and sentence comprehension subtasks, (Cronbach’s Alpha) ranged between 0.88 and 0.95. The data set for blind and deaf learners was based on national census and had similar reliability scores. This is what one of the interviewees had to say about the quality of the data used to set the benchmarks:

“The external evaluation data had this long passage and short passage so you could have comprehension items that are from a longer fluency story and a shorter fluency story. We had done the analysis in two directions. One, to get the mean scores, and another more critically to establish the levels of fluency, letter knowledge, and decoding knowledge, and know the listening comprehension knowledge was associated with reading comprehension. All we care about is comprehension. But we want to have an idea of what’s the level of other skills required to get to that comprehension. And ... what were the average fluency rates for the two different tools? What are the sentence fluency levels? So, it’s everything trying to ladder up to comprehension. What was the average fluency related to comprehension across these various tools?” Dr. Ben Piper.

The participants were divided into four groups: a) group one to work on draft benchmarks for English for regular Grade 1-3; b) group two to work on draft benchmarks for Kiswahili for regular Grade 1-3; c) group three to work on draft benchmarks for blind learners in

Grade 1 and 2; and d) group four to work on draft benchmarks for deaf learners for grade 1 and 2. Each group was provided with the relevant dataset.

For groups working on the benchmarks for regular learners, the data sets provided were based on the baseline and midline data collected by Management Science International (MSI) for *Tusome* which covered Grades 1 and 2. These datasets were representative nationally. For Grade 3, the teams worked with data sets from an Alternative Provision of Basic Education and Training (APBET) baseline study, which was representative samples from Nairobi, Kisumu, Nakuru, Eldoret and Mombasa cities – where APBET intervention was being undertaken. Groups 3 and 4 got datasets generated from the Special Needs Education baseline study undertaken nationally by *Tusome* in June 2017.

Using the data outputs, the teams were instructed to undertake the following tasks:

- a) Estimate low benchmark level
- b) Estimate high benchmark level
- c) Estimate the proportion of learners who reached each of these benchmark levels
- d) Estimate the proportion of learners who would reach the specified benchmark levels in 2022.

The RTI *Tusome* Monitoring and Evaluation lead provided technical support to the teams. Whenever teams made proposals on benchmarks, he would compute for them, the proportion of learners who reached the benchmark levels in the available data. This would often lead to heated debate and further deliberations on either increasing or reducing the initially proposed benchmark level.

After the discussions, when groups had agreed on the benchmarks, the lead for each of the four teams made a plenary presentation on benchmarks, providing reasons for the benchmarks they had proposed and sought further input from workshop participants. The workshop eventually agreed on the draft benchmarks as shown in Table 4.

Table 5: Draft 2019 Benchmarks

School type and grade	Subject and Level			
	English		Kiswahili	
	Low	High	Low	High
Regular- Grade 1 (cwpm) ⁴	20	35	10	30
Regular- Grade 2 (cwpm)	30	65	17	45

⁴ Cwpm stands for ‘Correct words per minute’, while clpm stands for ‘correct letters per minute’.

Regular- Grade 3 (cwpm)	40	80	30	55
Blind – Grade 1 (clpm)	10	15	10	15
Deaf – Grade 1 (clpm)	15	35		
Blind – Grade 2 (cwpm)	15	20	10	15
Deaf – Grade 2 (cwpm)	10	25		

(Development of Benchmarks on Early Grade Reading (Report availed by Dr. Ben Piper)

Further, teams presented the proportion of pupils expected to reach the benchmarks by 2022 as shown in Table 6.

Table 6: Percentage of pupils to reach the proposed benchmark by 2022

School type and grade	Subject and Level			
	English		Kiswahili	
	Low	High	Low	High
Regular- grade 1	35 %	25%	30%	12%
Regular – Grade 2	60%	25%	80%	24%
Regular- Grade 3	35%	20%	35%	10%
Blind – Grade 1	30%	28%	27%	23%
Deaf – Grade 1	50%	30%		
Blind – Grade 2	34%	30%	33%	25%
Deaf – Grade 2	30%	15%		

(Development of Benchmarks on Early Grade Reading (Report availed by Dr. Ben Piper)

4.1.3 Briefing to the KNEC CEO

The next step was a briefing meeting with the KNEC Chief Executive Officer. RTI *Tusome* was invited to make a presentation on the process of benchmark development to the KNEC CEO. The briefing was held at the CEOs office in March 2019 with 10 participants in

attendance (6 male and 4 female). In taking the CEO and the participants through the process of benchmark development, the presenter also covered deliberations during the process of developing benchmarks for Kenya, then shared the draft benchmarks developed during the stakeholder workshop. The KNEC CEO then requested that the draft benchmarks be presented to the KNEC steering committee for discussion and adoption.

4.1.4 Presenting draft benchmarks to the NASMLA Steering Committee Meeting

The draft literacy benchmarks were presented to the NASMLA steering committee in a meeting held in April 2019. The meeting was held at the KNEC Council Room and was attended by 18 participants (10 male and 8 female) drawn from various organizations as shown in Table 7.

Table 7: Steering Committee Members present in the meeting

Name of Institution	Designation of representative
Kenya National Union of Teachers	Assistant National Treasurer
University of Nairobi	Senior Lecturer
Kenya Institute of Special Education	Deputy Director
Kenya National Commission for UNESCO	Director, Education Programs
National Parents Association	National Chairman
RTI- <i>Tusome</i>	3 officers: Chief of Party; D- COP Programmes; SEPC (SNE); M&E Director;
Kenya National Examination Council	4 officers: CEO; Head of National Assessment Centre; Deputy Director; and Research Officer
Ministry of Education	Senior Assistant Director of Education
Twaweza East Africa	Country Manager
Kenya Private Schools Association	National Chairperson
MoE-Directorate of Policy & Development	Senior Assistant Director of Education
Women Educational Researchers of Kenya	Chairperson

(Development of Benchmarks on Early Grade Reading (Report availed by Dr. Ben Piper)

During the meeting, the RTI-*Tusome* team made a presentation on the benchmarking process. Though members of the steering committee were receptive, they noted that the presentation was quite technical and so they asked the then head of NAC to comment on the methodology of benchmark development, and she confirmed that the process had been undertaken with the required technical rigor and that she did not have any reservation with the process and the output. The committee particularly commended the development of benchmarks for blind and deaf learners as a milestone in Kenya.

In this meeting, the draft benchmarks were formally adopted as proposed and without adjustment.

In later *Tusome* documents, the benchmarks are presented across four categories of readers: Zero reader, beginning reader, emergent reader, and fluent reader as summarised in Table 8 below:

Table 8: Benchmarks for oral reading performance categories

CATEGORY	ENGLISH			KISWAHILI		
	Grade1	Grade 2	Grade 3	Grade1	Grad 2	Grade 3
Zero- readers	0	0	0	0	0	0
Beginning readers	1-19	1-29	1-39	1-9	1-16	1-29
Emergent readers	20-34	30-64	40-79	10-29	17-44	30-54
Fluent readers	35+	65+	80+	30+	45+	55+

(USAID, 2022, p.16)

4.2 The second set of Benchmarks

Information from the interviews indicated that, for early grade, the second set of benchmarks is provided for only Grade 3 English at 90 cwpm. This benchmark is referenced in the KEYA scoring guide and in the curriculum designs for Grade 3. The curriculum designs also reference 30 cwpm for Grade 1(which is different from the 35 cwpm as developed in the first set and 65 cwpm for Grade 2 the same as in the benchmark in the first set.

The development of these new national benchmarks was driven by the paradigm shift in the curriculum and the attendant pedagogical approaches. Though ways of categorising learner achievements in reading were already available in the NASMLA and the *Tusome* reports, the interviewees indicated that these reports were not accessible to teachers, and therefore, it was important to assess the new curriculum using the same language that the teacher is familiar with so that there is *“a marriage between what the teacher sees in the curriculum designs and what is seen in the assessment documents”* (KNEC officer).

Documentation of the process through which this second set of benchmarks was developed was not available in official or semi-official documentation. Therefore, the information presented here is based on interviews conducted with KNEC officers and other key informants. The perception was that the new curriculum needed to be assessed in a new way. Therefore, as part of preparing to assess the CBC curriculum, a meeting was called at Kiambu Institute of Science and Technology for the purpose of developing tools for monitoring learner achievement. According to the informant, the participants in this meeting were teacher trainers for languages, practising teachers for languages, the Kenya Institute of Curriculum Development, the Quality Assurance department in the Ministry of Education and KNEC. The process through which these benchmarks were set is captured in the quote below from one of the interviewees:

“We started with a process whereby we had to look at the documents that were available to us at that time. Of course, the ones not being available at that time... but just an experience from what the teachers had experienced because they (the earlier benchmarks) weren't accessible to us at that point in time. So, we looked at the international global benchmarks...we looked at what the curriculum says the range at which the learner should be able to be operating at ... in terms of the texts that they [Grade 3] are supposed to be reading. And then we looked at what it is that is meaningful for one minute at grade three. And that is how we set the benchmark. And we've been sustaining it all through the MLPs that we have done”. KNEC Officer.

Information from the scoring guide for KEYA 2024 indicates that for ORF the reading benchmark is categorised into four levels of competency as outlined in Table 9 below:

Table 9: KNEC Reading Benchmarks for Grade 3

Performance Level	Correct Words Per Minute (cwpm)
Below Expectation (Level 1)	0-44
Approaching Expectation (Level 2)	45-89
Meeting Expectation (Level 3)	90
Exceeding Expectation (Level 4)	90-157

(KNEC, Grade 3 SBA-2024 Scoring Guide)

Because the Grade 3 reading benchmark is available in the curriculum design and in the scoring guide for KEYA, and because KEYA is administered to all Grade 3 learners following the CBC curriculum in Kenya, this benchmark is nationally available to all teachers teaching lower primary. The curriculum designs are also available in the KICD website, and, therefore, stakeholders can access information on the benchmark cwpm for a Grade 3 in Kenya.



4.3 Recommendations for way forward on Kenyan literacy benchmarks

While one of the interviewees was categorical that there are two sets of benchmarks, and that the first one is 'programmatic' and, therefore, cannot be taken as the national benchmark, a close scrutiny reveals that the ORF benchmark used for KEYA and cited in the curriculum design is an offshoot of the benchmark setting process undertaken during the *Tusome* intervention. In fact, one of the interviewees from KICD indicated that what is given in the design and in the KEYA scoring guide, "***act as an example of how to graduate a rubric and, therefore, [we] cannot claim to have two sets of benchmarks***". Further, what is in the curriculum designs for Grade 2 is the same as was developed during *Tusome*, and what Grade one the curriculum design gives (30 cwpm) is not far from the *Tusome* 35 cwpm for English. Thus, the difference between the two is not big. Therefore, given the comprehensive process that developed the earlier benchmarks, and the current changes in curriculum, it is time for Kenya to review/revise the benchmarks using as rigorous a process as was used initially, and to have a clear plan for integrating these benchmarks into the teaching, learning and assessment processes in a more consistent way.




5. Lessons learned from the Kenyan context

A reflection on Kenya's drive to improve learner achievements in foundational literacy, to assess the progress in this engagement, and to set benchmarks that form a basis for evaluating progress, reveals that a structured, rigorous and systematic instruction programme, focusing on the development of decoding skills, has the potential to improve early grade literacy even in a complex multilingual setting like Kenya. There are a number of lessons that Kenya can offer to other countries, to the global community and perhaps to UIS for its engagement in the three areas.

5.1 Lessons from Kenya's drive to improve foundational literacy

Despite the challenges occasioned by the implementation of curriculum reforms in the middle of the *Tusome* intervention, and the impact of the loss of learning during the COVID-19, Kenya's up scaling of the gains of PRIMR through the *Tusome* nationwide intervention stands out as one of the most successful drives on foundational literacy. It provides evidence that focus on Kenya has learned a lot through both the success and the challenges. There are several lessons Kenya can share with other countries:

- a) One of the lessons is the need for when a government to establish or strengthen the office that coordinates donor funded projects to ensure the mainstreaming of the good practices that come with projects so that these practices can lead to policy level discussions that inform key decisions. Without such an office, the projects gains will die or weaken when the project funding comes to an end. For example, though *Tusome*, was a ministry of education scale up, the gains have been weakening because the pedagogical approaches did not get anchored in policy. Thus, many stakeholders look at *Tusome* approaches as programmatic, and CBC approaches as the Ministry of Education policy. This makes a big difference particularly to practising teachers.
- b) Related to the above is a lesson on teacher training and support. While it is true that in-service training and coaching support are powerful because they address the practical issues on the job, an early literacy intervention programme that hopes to have a lasting impact on classroom practice must incorporate modules for pre-service teacher programmes. In Kenya, much of the *Tusome* organisation of literacy content for both English and Kiswahili is part of the Ministry of Education curriculum designs for Grades 1 to 3, but this has not yet been integrated into the pre-service teacher training. This presents a challenge for continuity of the structured pedagogy even when the teachers' guides are available. A programme of this nature needs to have a module for pre-service teacher training.
- c) Another lesson relates to the multilingual context. When a majority of the learners join school without the language of instruction, then even with the most effective literacy interventions like *Tusome*, a big proportion of the learners will still be struggling with reading comprehension at the end of Grade 3 because their learning engagement is double: they are learning the language (oral vocabulary), and at the same time they are learning to read and write (literacy skills). This calls for the country to be deliberate in settling realistic targets based on local data; to




technically and not politically determine the time allocated to language instruction (for the language used as the medium of instruction) in early grades; and to institutionalise a ‘catch up’ or remedial instruction in Grade 3 to support struggling readers before, or, as they transition onto upper primary.

5.2 Lessons from Kenya’s measurement of learning progress

On the aspect of measurement of foundational learning, the umbrella lesson from Kenya is that a country can be very busy with many national assessment programs and still fail to report to UIS. Kenya has many sets of data on various aspects of learner achievement in early literacy, from the *Tusome* programme, from NASMLA Class 3, and more recently from KEYA (MLP Grade 3), but this information remains in the national reports and project reports, and yet this has not resulted in reporting to UIS. What specific lessons can we draw from this?

- a) The first lesson is on the importance for a country to have an agreement on what to measure, how to measure it, and how to report the results. As indicated above, Tusome reporting provides data on learner performance on the EGRA sub-tasks and on the national benchmarks developed collaboratively with the Ministry of Education while NASMLA reports using adapted SEACMEQ categories, and MLP came up with new categories and new benchmarks for Grade 3 aligned to CBC. The three are national assessments, but they are not comparable and according to the documents reviewed, and the information obtained from interviews, none of them quite qualifies for UIS reporting, but could do so with a bit more work.
- b) The second lesson is on the establishment of a national assessment center to coordinate measurement of learning outcomes at national level. The National Assessment Centre (NAC) in KNEC has the experience of conducting national assessments of learner achievement, and though the interviews revealed that there are some technical training gaps, the systems are in place. Kenya just needs to identify the nature of the specialised technical skills needed by the personnel in NAC so that they can improve one of the available national assessments to meet the global reporting rigor. International guidance from UIS, or from NGOs or others working within the UIS mandate or orbit and who understand the UIS criteria and standards, would most likely be highly welcome. Even online guidance from UIS through some sort of electronic portal would be useful.
- c) A third lesson on measurement relates to the need for a country to establish a clear path for reporting path to UIS. While it was clear from the interviews, who is mandated to conduct the national assessments, it did not seem clear to my interviewees which section of the Ministry of Education is mandated to do the reporting if the available data met the reporting standards.
- d) Finally, Kenya offers a lesson on the positive impact related to the use of (adapted) EGRA tools. An assessment using these tools provides detailed reports on learner progress in decoding skills. Assessment is not just about reporting; it is also about improving, and the EGRA tools give reports which are easy to convert into intervention. A recent study analysed reading assessment data of over half a




million pupils from 48 LMICs tested primarily in a language of instruction and concluded that “Literacy goals in Low- and Middle-Income Countries will be reached only by ensuring focus on decoding skills in early-grade readers. Effective literacy instruction will require rigorous systematic phonics programmes and assessments suitable for LMIC contexts” (Crawford, et al. 2024, par.1). UIS could consider what improvement could be made on EGRA tools for countries which are already familiar with their use, and which are struggling with a second/foreign language of instruction in a multilingual setting.

5.3 Lessons from Kenya’s benchmark setting experience

The benchmarking process in Kenya offers a number of lessons:


- a) The first one is that using rich local data when developing reading benchmarks provides benchmarks which are trusted and respected by the stakeholders. All the stakeholders interviewed for this study expressed acceptance of the benchmarks set with the support of RTI *Tusome* team because they are based on Kenyan data. The only complaint aired by one of the interviewees is that these benchmarks were aligned to the system that the country is phasing out.
- b) A second lesson from the Kenyan experience is that benchmark setting can be inclusive as long as there is adequate data set, and the process involves all relevant stakeholders and technical support. In one sitting, the Kenyan team was able to set benchmarks for multiple structurally unrelated languages, and even for learners with special needs.
- c) The Kenyan experience also provides a lesson on the need for external technical assistance even when a country has a vibrant national assessment centre. While staff from NAC steered the benchmarking workshop and the participants included experts representative of the stakeholders, the technical assistance was provided by the *Tusome*- RTI staff, who workshoped the team of experts on the process of benchmarking, provided comprehensively analysed and reported Kenyan data based on adapted EGRA sub-tasks, and supported the groups by providing statistical implications of the draft benchmarks that groups would propose during the process. In fact, one of the NAC staff indicated that his role in the benchmarking workshop was that of a ‘technical observer’. However, when the benchmark setting process is undertaken with technical assistance from a funded project, official adoption needs to be followed by a policy decision, to ensure that these benchmarks get institutionalised. Had there been a policy position after the adoption of the national benchmarks in 2019, these would have been the benchmarks to be reviewed/revised after
- d) Another lesson relates to embracing the SDGs. All the reports reviewed, and all the people interviewed cited a link between the efforts to improve and track the progress in early literacy and Kenya’s commitment to report on SDG 4.1.1a. Thus, the fact that the SDGs had been set, and that UIS had published guidelines on reporting, provided both inspiration and pressure in setting the benchmarks to help the country evaluate learning progress at this level.

- 
- e) Finally, the interviews indicated that the capacity of the staff in the NAC has improved tremendously through engagements with international community of scholars and technical experts working within the UIS frameworks. In 2021, USAID funded two, one-week policy-linking workshops in Kenya to set provisional global benchmarks on the grades 2 and 3 national learner assessments in English and Kiswahili. The Kenyan team assembled in one place, but the workshop was conducted with virtual lead facilitators. This workshop was preceded by a virtual one-week capacity building training for KNEC specialists. Though only one of the four assessments subjected to the policy linking process was found to be ‘strongly aligned’, the interviewees from NAC indicated that this engagement with the international community opened their eyes to the skills gaps in their center with relation to UIS reporting standards. Information from the interviews also indicated that some of the assessments have already been revised to cover the required sub-tracts, and that there are other capacity building engagements with the international community planned for early 2025. Other countries can leverage this kind of support.




REFERENCES

- Crawford, M., Raheel, N., Korochkina, M., & Rastle, K. (2024). *Inadequate foundational decoding skills constrain global literacy goals for pupils in low- and middle-income countries*. *Nature Human Behaviour*. <https://doi.org/10.1038/s41562-024-02028-x>.
- Kenya National Examinations Council (KNEC). (2010). *The report on monitoring learner achievement study for class 3 in literacy and numeracy*. KNEC.
- Kenya National Examinations Council (KNEC). (2016). *Monitoring learner achievement study for class 3 in literacy and numeracy in Kenya*. KNEC.
- Kenya National Examinations Council (KNEC). (2019). *Report on monitoring learners' progress Grade 3*. KNEC.
- Kenya National Examinations Council (KNEC). (2020). *Monitoring learner achievement at class 3 level of primary school education in Kenya*. KNEC.
- Kenya National Examinations Council (KNEC). (2022). *Report on monitoring learner progress at Grade 3*. KNEC.
- Mugo, J., Kaburu, A., Limboro, C., & Kimutai, A. (2011). *Are our children learning? Annual learning assessment report*. Uwezo.
- Piper, B., Zuilkowski, S., & Mugenda, A. (2014). *Improving reading outcomes in Kenya: First-year effects of the PRIMR Initiative*. *International Journal of Educational Development*, 37, 11–21. <https://doi.org/10.1016/j.ijedudev.2014.02.006>.
- Piper, B., & Mugenda, A. (2012). *The Primary Math and Reading (PRIMR) Initiative: Baseline report*. Prepared under the USAID Education Data for Decision Making (EdData II) Project, Task Order Contract No. AID-623-M-11-00001. RTI International.
- UIS & ACER. (2022). *COVID-19 in Sub-Saharan Africa: Monitoring impacts on learning outcomes. Main report*. UNESCO. http://milo.uis.unesco.org/wp-content/uploads/sites/17/2022/01/MILO-Main-ReportSSA-Jan-2022_EN.pdf.
- UNESCO. (2024). *Assessments for minimum proficiency levels a and b (AMPL-ab) international report*. UNESCO Institute of Statistics.
- United States Agency for International Development (USAID). (2012). *The Primary Math and Reading (PRIMR) Initiative: Baseline report*. EdData II Technical and Managerial Assistance, Task Number 13.
- United States Agency for International Development (USAID). (2014). *USAID/Kenya Primary Math and Reading (PRIMR) Initiative: Final report*. Award No. AID-623-M-11-00001.
- United States Agency for International Development (USAID). (2015/16). *Tusome revised baseline study*. Task Order No. AID-623-TO-14-00004.
- United States Agency for International Development (USAID). (2017). *Tusome external evaluation – Midline report*. Contracted under AID-615-TO-16-00012.



United States Agency for International Development (USAID). (2020). *Tusome external evaluation – Endline report*. Prepared under Contract No.: GS-10F-0033M/AID-OAA-M-13-00010, Tasking N 7617.018.01.

United States Agency for International Development (USAID). (2022). *USAID Kenya Tusome early grade reading activity study on grade 1, 2, and 3 pupil learning outcomes in 2021: Final report*. Award No. AID-615-C-14-00007.



**Benchmark Setting for Precursor Skills in Reading:
A Case Study of Kenya**

Email:

uis.information@unesco.org

uis.director@unesco.org

uis.unesco.org

[@UNESCOstat](https://twitter.com/UNESCOstat)