Girls' Education South Sudan (GESS): Learning Assessment Endline

Funded by: DFID

Document Created: August 2018

Knowledge, Evidence and Research

The Knowledge, Evidence and Research (KER) component of the Girls' Education South Sudan (GESS) programme aims to generate increased knowledge and evidence for policymakers of what works to promote girls' education in South Sudan, about programmatic causality and impact, and to provide evidence, lessons learned to inform future programmes and other contexts. The KER programme develops an evidence base for the project interventions, linking inputs to outcomes and impacts, and gathers broader information about what works in girls' education. The Programme gathers data continuously through the South Sudan Schools' Attendance Monitoring System (SSSAMS), twice yearly through Longitudinal Qualitative Survey (LQS), yearly through the School Sample Survey, and then has set piece Baseline (2014), Midline (2016), and Endline (2018) survey waves.

Acknowledgements

The team at Charlie Goldsmith Associates (CGA) would like to thank our colleagues at the Ministry of General Education and Instruction (MoGEI), Mott MacDonald, the GESS State Anchors and education partners for their assistance in facilitating and conducting the Endline Learning Assessment and producing this report.

For their help in facilitating the research, including input into questions, facilitating access, accompanying researchers and valuable discussion of results and recommendations, we would particularly like to thank MoGEI, including, but not limited to:

Rt. Hon. Deng Deng Hoc Yai, Honourable Minister of General Education and Instruction

Michael Lopuke Lotyam, Honourable Undersecretary

George Mogga Benjamin

Victor Dut Chol

George Ali Steven

Clement Kajokole

Mading Manyok

Kur Ayai

Kiri Lokuto

GESS/CGA would especially like to thank colleagues at MoGEI who assisted with the data collection and field research. Specifically:

Clement Kajokole

Ayach Awan Thon

Saida Adam

Mathew Mangar Thon

Rubaya Rose Wani

Malish Kennedy Peter

Mabor Tur

Nuul Gabriel

Sarafino Tisa

Clement David

Alice Taddeo Bidai

Christopher Awii

Anguei Malei Majok

Rebecca Malok

Angelo Machar Akec

Deng Simon

Wudu William

Yom Mach Elizabeth

Taban Kozo

Doru Joyce

John Deng Kuer

We would also like to thank the GESS State Anchor teams for their invaluable logistical support to research teams in the field:

ADRA; Food for the Hungry; HARD; Stromme Foundation; UMCOR; and Windle Trust.

Finally, we would like to thank UK aid and our partners on the GESS programme for input into the tools, ongoing support and fruitful discussion of findings and recommendations:

Mott MacDonald/Cambridge Education; BBC Media Action; and Winrock International.

Contents

List of Tables

List of Abbreviations

BoG	Board of Governors
CED	County Education Department
CES	(former) Central Equatoria State
CG	Capitation Grant
CT	Cash Transfer
DFID	UK Department for International Development
EES	(former) Eastern Equatoria State
ETMC	Education Transfers Monitoring Committee
GESS	Girls' Education South Sudan programme
GRSS	Government of the Republic of South Sudan
GUN	Greater Upper Nile
IDP	Internally Displaced Person
JGL	(former) Jonglei State
KER	Knowledge, Evidence and Research
LKS	(former) Lakes State
MoGEI	Ministry of General Education and Instruction
NBG	(former) Northern Bahr el Ghazal State
PEO	Payam Education Office
PES	Payam Education Supervisor
PLE	Primary Leaving Examination
SCE	Secondary Certificate Examination
SMC	School Management Committee
SMoGEI	State Ministry of General Education and Instruction
SSSAMS	South Sudan Schools Attendance Monitoring System
UNS	(former) Upper Nile State
UTY	(former) Unity State
WBG	(former) Western Bahr el Ghazal State
WES	(former) Western Equatoria State
WRP	(former) Warrap State

Executive Summary

As a result of decades of civil war, South Sudan today faces huge challenges in developing its low educational base. The GESS programme aims to transform a generation of South Sudanese girls by increasing access to quality education. One of the strategic objectives of MoGEI is to eliminate barriers to girls' education and promote gender equality throughout the education system.

This Endline report presents the findings of the Learning Assessments conducted in a sample of schools from all ten former States.[i]

The 2018 GESS Endline Learning Assessments were carried out in all ten (former) States of South Sudan. The aligns with the 2016 Midline design, but is unlike the 2014 Baseline of 2014, which was carried out only in the seven States accessible at the time of study. Literacy and numeracy tests were administered in May and June 2018 at Primary 5 (P5), Primary 8 (P8) and Secondary 2 (S2) grades in a sample of 64 schools, including 44 primary schools and 20 secondary schools. This sample included assessments of 10,266 pupils, of whom 4,701 were girls and 5,565 were boys.

In 2014, results indicated a significant gender gap at all grade levels in both literacy and numeracy. While girls' scoring on the Learning Assessments remains lower than that of their male counterparts in all grades, girls narrowed the attainment gap from 2.5% in 2014 to 1.7% in 2018; however this represents a rising gap since Midline, when girls' and boys' average scores were almost equal. The sample is too small to draw conclusions as to whether there is a broader trend of the attainment gap widening since 2016.

As in 2014, pupils perform 'better' (i.e. receiving higher mean scores) on the literacy compared to the numeracy assessments, though results are lower than 2014 in four assessments. While there are challenges and limitations in evaluating the impact of a national programme, the results of the Baseline, Midline and Endline learning assessments contribute to the KER strand of GESS.

Results and Analysis

- Overall, girls' average scores decreased to 42.1% in 2018, having been 45.7% in 2016 and 44.2% in 2014. This result is likely influenced by the examinations taking place three months earlier in the school year than was the case for the Midline and

[i] The sample design, fieldwork, and analysis of the 2014 Baseline were conducted prior to the Executive Order in October 2015 establishing 28 States in South Sudan, and the Executive Order in December 2016 establishing 32 States. State-level findings in this 2018 report refer to the previous administrative delineation of ten states for consistency.

Baseline. Students in 2018 therefore had received a third less of the nine-month school year at the point of sitting the exam.[ii]

- In 2014, girls made up 32% of the sample across 48 schools. In 2016 girls made up 40% of the sample across 59 schools, and by 2018 the percentage of girls had increased to 48% across 64 schools. These percentages correspond with the gender ratio in the schools surveyed and is slightly better than the national enrolment gender ratio (approximately 46% of national students are girls).
- 2014 results indicated a gender gap at all grade levels in both literacy and numeracy. At the Endline there is still a gender gap at each level, but the gap has narrowed in four of the six assessments.
- Scores significantly increased from the Midline at S2 level. Overall S2 literacy scores went from 32.6% in 2016 to 48.6% in 2018. S2 numeracy scores went from 26.9% in 2016 to 37.2% in 2018. There has been a large increase in enrolment over the course of GESS, and improved results may be indicative of the students in S2 in 2018 having received more days of primary education than S2 students in 2016.
- Scores also increased from the Baseline in P5 numeracy and S2 numeracy.
- As well as the assessments taking place three months earlier in the academic year than previously, other reasons for the weaker results in 2014 include:
    - 2018 has seen significantly higher enrolment numbers. Students sitting assessments in 2014 and 2016 were therefore more likely to be from families who prioritised education, therefore achieving better results.
    - The difficult economic situation in 2018 may also have caused distractions for teachers and families, with pupils potentially spending less time in schools.
    - Continued insecurity in some areas has resulted in disrupted schooling due to displacement, and this may also have contributed to the 2018 fall in scores.
- Overall, and in line with findings from other KER activities, it appears that GESS interventions with respect to enrolment have had more impact than ones with respect to quality; this is consistent with the programme's approach whereby quality education activities were pilots in small numbers of schools, whereas activities impacting enrolment were implemented nationwide.

Key Recommendations

- Support needs to be given to pupils coming into primary education who may have missed some early school years. Remedial plans and materials should be put in place to support these students as they join or re-join education, while progression for existing students should not be impeded.
- Reported enrolment has reached record highs each year that GESS has been operating, from 0.9 million in 2014 to over 1.7 million in 2018. However, the results of the Learning Assessment show that average pupil attainment has declined in many of the papers. As access to education is broadened and in context of these results, it

[ii] The Learning Assessment test were administered earlier than Baseline and Midline to enable completion of analysis and reporting before the end of the Programme by 30th September 2018

is vital for future programmes to check that fundamental skills are being learnt at all levels of schooling, rather than starting to assess learning at P5, and focusing on the curriculum. This approach would protect the assessment from being impacted by the time of year it is conducted.

- More attention is required to reading and interpreting informational text at all levels of literacy instruction. Classroom materials should be developed that expose pupils to a range of different question types when interrogating texts. General vocabulary work in the classroom needs to be improved, e.g. the use of figurative language, which was an area in which pupils struggled.
- At primary level, there should be increased classroom time for studying the practical applications of mathematics and applying mathematical skills to real-life situations. At secondary level, more focus is needed on interpreting numerical data and spotting numerical patterns.
- If Learning Assessments continue under GESS2 and remain linked to the curriculum, they will need to be redesigned in light of the roll-out of the new South Sudanese curriculum in future.
- Collaboration with the Norwegian teacher training programme would be a realistic way of improving the quality of literacy and numeracy teaching in South Sudan.

1 Background

The GESS programme in South Sudan

The Girls' Education South Sudan (GESS) programme seeks to transform the lives of a generation of children in South Sudan – especially girls – through education.

GESS is an initiative of MoGEI and funded by UK aid. MoGEI leads the GESS programme, supported by implementing partners who provide technical advice. At State and County level the State Ministries of General Education and Instruction (SMoGEI) take the lead in programme implementation, supported by partner NGOs, or 'GESS State Anchors'. Implementing partners include Mott MacDonald/Cambridge Education (lead), BBC Media Action, Charlie Goldsmith Associates and Winrock International.

GESS is a practical programme that implements activities that tackle financial, cultural and quality barriers to education for the girl child, while boys will also benefit from an improved learning environment.

The activities are structured along three main outputs:

1. Enhanced household and community awareness and empowerment for supporting girls' education through radio programmes and community outreach.
2. Effective partnerships between the Government of the Republic of South Sudan (GRSS) and local organisations to deliver a community-based school improvement programme which will include:
    a. Cash Transfers to girls and their families;
    b. Capitation Grants to schools;
    c. Provision of practical support to schools, teachers and education managers to improve the quality of education.
3. Knowledge, Evidence and Research (KER) - increased knowledge and evidence of what works to promote girls' education in South Sudan.

The GESS programme was designed in 2012, shortly after South Sudan gained independence, and was officially launched in April 2013. The Programme is monitored and evaluated on the basis of several tools, including the Learning Assessment.

In 2016 and 2018, the GESS Learning Assessments were administered across all ten former States of South Sudan (in 2014 assessments were only administered in the seven States participating in the programme at that point[iii]). The Learning Assessments are confidential materials, as they may again be used as 'live' tests in future years; therefore, test security protocols apply. Mindful of this, specific questions and text content for the literacy

[iii] Central Equatoria, Eastern Equatoria, Lakes, Northern Bahr el Ghazal, Warrap, Western Bahr el Ghazal and Western Equatoria

assessments are not included in this report. These can be viewed separately on request to GESS.

Background to the Learning Assessment component of GESS

The GESS Learning Assessments were originally developed and piloted in South Sudan by Education for Change (EfC) in 2013. The findings of these trials were incorporated into the 2014 Learning Assessments. In July 2014, Forcier Consulting, a specialist research consultancy working in South Sudan, was appointed to review the Learning Assessment materials, to train invigilators and markers, to manage the marking and coding of pupils' responses, to analyse the data and to provide a report on outcomes and recommendations. The 2014 assessment served as a Baseline against which the 2016 Midline and 2018 Endline results are compared. The contribution of all earlier contributors is hereby acknowledged and appreciated.

An update on the South Sudanese context since the Baseline and Midline Reports

The security and economic situation in South Sudan has deteriorated significantly since the Baseline. Despite this, enrolment has increased by 800,000 over the last five years; however, up to 2.4 million South Sudanese children are still not in schools within the country.[iv]

The context in which GESS operates has deteriorated since 2014, as the dynamics of the conflict that broke out in December 2013 have shifted, and the economy has collapsed. Fighting was initially concentrated in the Greater Upper Nile region, but later spread to areas that had previously been relatively stable, in particular the Equatorias and former Western Bahr el Ghazal (WBG) State. This resulted in mass displacement both within South Sudan and into bordering countries, making it increasingly difficult for schools to function. The implementation of the Agreement on the Resolution of the Conflict in South Sudan (ARCISS), signed by both parties to the conflict in August 2015, has faced numerous obstacles, the most significant being the fighting that erupted in Juba in July 2016, and the intensification of the conflict elsewhere. In July 2018, the number of South Sudanese seeking refuge in neighbouring countries stood at 2.47 million, with an additional 1.74 million Internally Displaced Persons (IDPs).[v]

This deterioration in the security and humanitarian situation has taken place against a backdrop of economic collapse. The fall in oil production has severely eroded the Government of South Sudan's chief source of revenue; meanwhile inflation has accelerated, with the effective South Sudanese Pound (SSP):USD exchange rate increasing

[iv] https://www.unicef.org/southsudan/media_21715.html
[v] Figures taken from OCHA's South Sudan Humanitarian Bulleting, July 15 2018, available at: https://reliefweb.int/sites/reliefweb.int/files/resources/20180716_OCHA_SouthSudan_Humanitarian_Bulletin%236.pdf

from 4.61:1 in September 2014 to 76:1 in September 2016, and beyond 300:1 in early 2018. This has had a strongly negative impact on education delivery as a whole, eroding the value of teachers' salaries and affecting the ability of schools to execute their budgets.

While the education sector has continued to operate, its ability to do so effectively has been hampered by the challenging security and economic context. As at August 2018, there were over 4,000 schools open in South Sudan, with 1,705,433 pupils enrolled, taught by around 40,000 full-time teachers, according to data on the SAMS (www.sssams.org - the near real-time management information system developed as part of GESS). However, UNESCO estimates that between 2.2 and 2.4 million children are still out of school in South Sudan, a number that is likely to rise.[vi]

A new peace agreement was signed in August 2018 and oil production is set to start again with the support of Sudan. It remains to be seen how the context of the country will change as the GESS programme comes to an end and GESS2 begins.

[vi] Global Initiative on Out of School Children: South Sudan Country Study,UNESCO 2018, available at: http://unesdoc.unesco.org/images/0026/002653/265399e.pdf

2. Purpose of Survey & Methodology

2.1 GESS Knowledge, Evidence & Research Objectives

The KER sub-output of the GESS programme seeks to generate knowledge and evidence about education in South Sudan, and what works to get girls in school, staying in school, and learning in school.

The research is focused on:

- Whether the Programme is achieving expected outcomes
- How outcomes are being achieved
- Wider areas of interest about what's happening in schools

The overall GESS research is based on the following two overarching questions, which have been developed from the outcome of the Programme:

- Has there been a change in enrolment and retention for girls and boys from Primary 5 to Primary 8 and from Senior 1 to Senior 4, and which aspects of the Programme contributed towards this?
- Has there been a change in the quality of education, as demonstrated by improved learning for Primary 5 to Primary 8 and Senior 1 to Senior 4? What changes in the learning and teaching environment have contributed to this?

The overall objectives of the GESS project surveys are:

- To monitor changes currently occurring in schools, particularly changes related to the GESS programme;
- To identify aspects of the GESS programme contributing towards changes in the enrolment rate among girls and boys Primary 5 to Primary 8 and Senior 1 to Senior 4;
- To identify aspects of the GESS programme that will contribute toward the future measurement of girls' and boys' retention rates between Primary 5 and Primary 8 and Senior 1 and Senior 4.

The overall KER component of GESS seeks to:

- Develop National and State capacity for research and use of evidence;
- Develop knowledge about the impact of project interventions;
- Develop broader information about what works in girls' education;
- Incorporate process monitoring into learning about successes and failures in
- design and implementation, protect against doing harm and monitor value for
- money;
- Inform policymaking: budget priorities and targeted support.

The Programme outcomes are directly concerned with improvements in enrolment, retention, and learning. Alongside the Learning Assessment, three other areas of research

were developed to enquire more in-depth information about relationships, activities, and processes linking programme interventions to the outcomes were proposed. These are school and classroom practices, educational choices by households and girls, and management capacity and structures.

Complementary to the Learning Assessment:

- A detailed School Sample Survey, incorporating interviews with learners, Head Teachers, teachers and representatives of school governing bodies in addition to lesson observations and building assessments was carried out in June – September 2018. The purpose of the survey was to build a picture of the state of schools in South Sudan and understand the educational experiences of pupils – in particular girls – teachers, and managers. The survey, which incorporated questions from the pilot School Sample Survey, also looked at the impact and effectiveness of Capitation Grants (CGs) and Cash Transfers (CTs), as well as the use of Daily Attendance Registers (DARs)
- In-depth Household Surveys were conducted throughout June - September 2018, using a subset of schools selected for the School Sample Survey to obtain a detailed picture of the sensitive and complex nature of household decisions about money, gender dynamics and power structures, as well as the experience of pupils and their households in and out of school. The Household Report provides contextual background that will help inform future changes in education patterns by providing details of household level decisions that affect enrolment and retention of girls in schools.
- A County and Payam Education Managers Survey was conducted in the same timeframe, aiming to collect the views and experiences of County and Payam education staff, and the ways their work may have been impacted by GESS.

2.2 Purpose of the Learning Assessments

The principal purpose of this Learning Assessment is to enable GESS to evaluate the success of the Programme. The 2014 iterations of the assessments acted as the Baseline against which to evaluate the effect of the interventions made. The 2016 Midline Learning Assessment and this 2018 Endline Learning Assessment enable the comparison of pupils' performance over time, particularly in terms of encouraging girls into school and demonstrating an increase in learning outcomes.

At a national level, the fact that this Assessment has taken place for the purpose of Programme evaluation should not undermine the validity of the assessment. However, the sample size is not sufficient at State level to use the statistical data for system monitoring on a State by State basis, nor is it robust enough to report at a school level. State-level data is therefore not discussed in the body of this report; however, selected State-specific and school-specific groupings have been included as an additional check on the aggregate results, but are not intended to be considered as discrete outcomes. Moreover, the comparability of pupils' performance over time is intended to evaluate the performance of

the cohort, not of individual pupils. The need to include a sample of urban and rural schools may skew the sampling process. Remarks on test validity can be found in the 2014 Baseline report.

The curriculum in South Sudan has evolved over time, including over the period of the GESS programme. Previously, some areas of the country continued to teach Sudanese or Ugandan curricula, but this variance was eliminated following a 2015 MoGEI decision to phase out all foreign curricula. This historical variance may have impacted on learning outcomes in different areas.

2.3 Methodology

The monitoring and evaluation of the GESS programme includes a summative assessment of learning, which requires that pupils' learning be measured at a set point in the school year. This consistent timing is particularly important for year-on-year comparisons. Unfortunately, the close of the first GESS programme in September 2018 necessitated the Endline being conducted in May/June 2018, rather than August/September as per the Baseline and Midline. This resultant loss of three months of student learning from the assessment must be considered when evaluating the 2018 Learning Assessment results.

For the GESS programme, the Learning Assessment 'standard' has been set based on the literacy and numeracy curriculum at P5, P8 and S2 grade levels. There are six Learning Assessments in total, designed to identify pupils' understanding of literacy and numeracy at these three grade levels. The layout of the examination has two configurations for each assessment in order to reduce cheating by copying from neighbours in the exam. The question content remained the same as the 2014 and 2016 assessments. A description of the review of the 2013 pilot materials, test development, and formatting can be found in the 2014 Baseline report.

2.3.1 Test collection

TABLE 1 TEST ADMINISTRATION DATES

State	Field date
Central Equatoria	w/c[vii] 21st May
Eastern Equatoria	w/c 21st May
Jonglei	w/c 4th June
Lakes	w/c 28th May
Northern Bahr el Ghazal	w/c 25th June
Unity	w/c 4th June
Upper Nile	w/c 4th June
Warrap	w/c 14th May
Western Bahr el Ghazal	w/c 14th May

[vii] w/c=week commencing

Western Equatoria	w/c 14th May

GESS State Anchors [viii] were responsible for coordinating with the schools that were assessed and for disseminating exam papers. The State Anchors were also responsible for collecting the exam papers (including unused papers) and bringing them to Juba. Invigilator reports and attendance sheets were collected from schools, which confirmed the conditions under which the assessments were conducted and enabled reconciliation between the number of students who had signed the attendance sheet and the number of assessments received in the CGA Juba office. No issues were noted with the invigilator reports and attendance sheets that were received.

As the Learning Assessment model is not designed to report on the performance of individual pupils, the papers are not 'marked', but rather a team of markers are trained to 'code' pupils' responses based on the answer options, unclear intent, or omissions. Staff conducted specialised training in both data entry and coding.

Data entry and quality control training occurred on 4th June 2018 for ten staff members. This training included 'standardisation training to ensure consistency with coding decisions. Staff not available for training on this date were individually trained at a later date prior to starting work. Tests were coded as they arrived, including a quality assurance exercise on the first day of 'live' coding. Data entry tracking sheets were integrated throughout the coding process. Overall coding and data entry occurred between 4th June and 4th July 2018. The coding process remained similar to the procedures in 2014 and 2016.

Where a student did not answer a question the data entry team entered 'No answer given'. If a student gave two answers to a question that only required one, the data entry team also entered 'No answer given'. Unclear answers were reviewed by two coders. If the two coders could not agree on the answer provided by the student, 'No answer given' was entered for that question. This coding system ensured students could not receive credit for a question unless they demonstrated clearly that they knew the answer.

TABLE 2 MAXIMUM MARKS PER LEARNING ASSESSMENT

Learning assessment (maximum # marks)		
Grade	**Literacy**	**Numeracy**
P5	16	30
P8	19	32
S2	30	32

2.3.2 The 2018 School Sample

Schools were the primary unit from which the sample was selected, and then within every school relevant classes were tested. The sample design stipulated that five primary and two

[viii] NGO partner organisations that implement the GESS programme in each of the ten former States.

secondary schools be selected per State for a total of 70 schools. However, some schools were inaccessible or not open at the time of fieldwork and were not replaced or substituted. In former Upper Nile State (UNS) the invigilator visited eight schools rather than seven. After the final data cleaning and quality control checks, a total of 64 schools have been included in the 2018 sample. The final sample size is ample for aggregate (national) analysis, but is not intended to be representative at a State or school level.

TABLE 3 2018 SAMPLE - SCHOOLS PER STATE

Former State	Primary		Secondary	
	Planned	Actual	Planned	Actual
Central Equatoria	5	5	2	2
Eastern Equatoria	5	5	2	2
Jonglei	5	4	2	2
Lakes	5	5	2	2
Northern Bahr el Ghazal	5	3	2	2
Unity	3	3	2	2
Upper Nile	4	5	2	3
Warrap	5	5	2	2
Western Bahr el Ghazal	5	5	2	2
Western Equatoria	5	4	1	1

The number of pupils at each grade level in the 2018 sample is similar to the 2016 sample. The Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ) recommendations state that roughly 2000 pupils should participate in each test to allow for robust statistical analysis and comparison.[ix] As the Learning Assessment sample was based on schools and not on individual pupils, the total number of pupils tested was dependent on enrolment at the schools selected for the sample. No further sampling was conducted by grade level. In practice, this means that the number of pupils in the sample reveals the reality of primary and secondary enrolment in South Sudan, particularly in reflecting the fact that school enrolment generally decreases at each higher grade level. Students leave education for a variety of reasons, including pursuing secondary education abroad, typically in Uganda or Kenya. In other circumstances, education is not financially viable for families, or social dynamics discourage girls from pursuing further education.[x]

[ix] SACMEQ is a collaborative network of 15 Ministries of Education that conducts educational policy research.

[x] This matter has been expanded upon in other GESS literature and has motivated the Cash Transfer and Capitation Grant elements of the Programme.

TABLE 4 2018 SAMPLE - GENDER RATIOS BY LEARNING ASSESSMENT

Grade	Overall Total	Girls Total	%	Boys Total	%
Literacy					
P5 Literacy	2363	1201	51%	1162	49%
P8 Literacy	1493	652	44%	841	56%
S2 Literacy	1394	552	40%	842	60%
Numeracy					
P5 Numeracy	2223	1119	50%	1104	50%
P8 Numeracy	1473	662	45%	811	55%
S2 Numeracy	1320	515	39%	805	61%

As with grade level, no further sampling was done to ensure equal gender representation. The ratio of boys to girls presented in this report therefore simply reflects who was in the classes tested, which tends to match the national enrolment pattern. Overall, girls make up 46% of the 2018 sample, which happens to be the same ratio as the national proportion of female students. The 2018 gender ratio has improved from the 2016 sample, in which girls accounted for 40% of the sample, and further from 2014, in which girls accounted for 32% of the sample.

In previous years the ability to draw firm conclusions about the gender ratios present in each grade was limited by a large number of instances where no gender was listed. In 2016 there were several grade levels with high rates of gender omission, as high as 14.48% of the sample in the case of P5 numeracy. In 2018, all exam papers asked students to state whether they were male or female. In the data entry coding, coders were not able to say 'No sex given', so all papers had to be assigned to a male or female. In the rare case of a student not stating their gender, coders were trained to deduce the student's gender based on the name they provided on the paper. Invigilator attendance reports were also used to establish the gender of students.

3. Findings

The differences in the sample frame must be considered when analysing results at the aggregate (national) level. The sample in 2014 included seven out of ten possible States, while the 2016 and 2018 samples include schools from all ten (former) States. While the larger sampling frame still does not allow for comparison at the State level, it is more robust at the national level than the 2014 sample. However, it does raise the possibility that 2014 scores may have been different had the sample included the inaccessible states. In particular, the Greater Upper Nile (GUN) States - former Upper Nile, Unity and Jonglei -

which are among the most conflict-affected and thus likely to have weaker education systems, were excluded from the 2014 sample. It is possible therefore that the 2014 results may have been lower in the aggregate had scores from these areas been included. State-level scores have been included in the Annex as a point of reference.

3.1 Assessment Performance

At the aggregate level, average test scores decreased from 45.7% in 2014 to 44.7% in 2016 and 43.0% in 2018. However, when disaggregated by assessment type, the average 2018 numeracy scores of 38.5% represent an increase from the 2014 Baseline of 35.8%, albeit this has decreased from 40.23% in 2016. Average literacy scores, however, had decreased from 55.5% in 2014 to 47.3% in 2018.

TABLE 5 AVERAGE SCORES BY SEX

Grade	Year	Overall	Girls	Boys
		%	%	%
Literacy				
P5 Literacy	2014	54.7%	51.6%	57.2%
P5 Literacy	2016	48.8%	47.6%	50.3%
P5 Literacy	2018	42.3%	41.2%	43.3%
P8 Literacy	2014	62.0%	61.5%	62.3%
P8 Literacy	2016	63.5%	62.4%	64.4%
P8 Literacy	2018	54.1%	53.1%	54.8%
S2 Literacy	2014	52.8%	50.5%	53.5%
S2 Literacy	2016	32.6%	30.5%	33.7%
S2 Literacy	2018	48.6%	47.5%	49.4%
Numeracy				
P5 Numeracy	2014	40.5%	37.5%	43.5%
P5 Numeracy	2016	47.1%	48.6%	49.8%
P5 Numeracy	2018	41.2%	41.2%	41.2%
P8 Numeracy	2014	39.1%	37.2%	40.3%
P8 Numeracy	2016	41.8%	43.4%	41.7%
P8 Numeracy	2018	35.5%	35.3%	35.7%
S2 Numeracy	2014	30.2%	27.7%	30.9%
S2 Numeracy	2016	26.9%	27.5%	26.9%
S2 Numeracy	2018	37.2%	35.2%	38.5%

When girls' and boys' scores are considered in the aggregate, scores have decreased in every assessment other than P5 numeracy and S2 numeracy. However, the gender gap between boys and girls has closed since 2014, with the numeracy gap now at 0.6 percentage points (down from 2.1 percentage points in 2014), and the literacy gap now at 2.6 percentage points (down from 3 percentage points in 2014).

TABLE 6 AGGREGATE AVERAGE SCORES BY SEX

Year	Girls average score	Boys average score

2014	44.2%	46.7%
2016	45.7%	45.3%
2018	42.1%	43.7%

3.2 Numeracy

3.2.1 Numeracy Curriculum

The South Sudanese numeracy curriculum can be divided into three areas:

- Numbers and operations
- Measurement and geometry
- Information handling

These areas test the following abilities:

4. Conceptual understanding is essential and includes knowledge of numbers and understanding of mathematical vocabulary and symbols. The pupils need the ability to reason, produce examples and manipulate ideas.
4. Procedural knowledge requires pupils to select, use and apply mathematical processes in different situations and should be able to demonstrate this through examples.
4. Problem solving requires pupils to use their conceptual and procedural knowledge of mathematics in new situations.

The numeracy Learning Assessments were designed to reflect this curriculum.

3.2.2 Numeracy Results

The overall gap between boys' and girls' scores across the numeracy Learning Assessments is now only 0.6%, having been 2.1% at the Baseline. This indicates that the gender attainment gap for numeracy has narrowed over the course of GESS, with girls' attainment now much closer to that of boys.

In 2016 there was a significant increase in girls' scores at P5. Scores for both girls and boys at P5 dropped in 2018, but while in 2014 the girls' scores were 6.1 percentage points lower than that of the boys, by 2018 their scores had evened out. At the Midline, girls' P8 were higher than boys' scores. However at the Endline girls' scores were less than those of the boys again. The gap between the girls' scores and the boys' scores narrowed to 0.5 percentage points in 2018 from 3.1 percentage points in 2014, although girls had been 1.8 percentage points higher than boys in 2016. S2 scores increased significantly from both the Baseline and the Midline, although having achieved better results than boys in 2016, girls slipped behind boys again in 2018.

At P5 and P8, scores were highest for questions involving operations on integers. Questions involving fractions scored much less highly. At S2, questions involving data interpretation and sequences were the most poorly answered.

3.2.2.1 P5 Numeracy

P5 numeracy assessments showed an increase from 2014 to 2016, but fell in 2018, with girls' results remaining above the Baseline results, but boys' dropping below. This represents an evening out of girls' and boys' results by 2018: in 2014 the girls' scores were 6 percentage points lower than that of the boys, but by 2018 the scores were the same.

Questions were included from across the P5 numeracy curriculum. They were all multiple choice. Approximately one-third were questions around numbers and operations, which increased in difficulty from two-digit questions to three digits and decimals. Other questions included fractions, measurements, problem solving and geometry.

Students found straightforward number operation questions involving addition or subtraction of whole numbers the easiest to solve. Questions that were in the form of written problems were less well answered, although again, addition or subtraction questions received the most correct answers.

TABLE 7 AVERAGE SCORES IN PRIMARY 5 NUMERACY

Year (P5 Numeracy)	Girls' average score	Boys' average score
2014	37.5%	43.5%
2016	48.6%	49.8%
2018	41.2%	41.2%

3.2.2.2 P8 Numeracy

The increase in girls' scores from 2014 to 2016 was not maintained in 2018 and girls' scores narrowly fell behind that of the boys again. However, the gap between the girls' scores and the boys' scores narrowed from 3.12% in 2014 to 0.3% in 2018.

The P8 Assessment included a greater range of questions from across the curriculum than the P5 Assessment, with fewer than 13% straightforward number operation questions. Other questions included ones on probability, ratios and graphs. 13% of the questions involved problem solving, and nearly 20% were on geometry.

Number operations remained the highest scorers, especially addition and subtraction. As in P5, questions about fractions resulted in low scores.

TABLE 8 AVERAGE SCORES IN PRIMARY 8 NUMERACY

Year (P8 Numeracy)	Girls' average score	Boys' average score
2014	37.2%	40.3%
2016	43.4%	41.7%

2018	35.3%	35.7%

3.2.2.3 S2 Numeracy

S2 numeracy assessments demonstrated the breadth of the mathematics curriculum.

In 2014 the average girls' scores was 3.2 percentage points lower than that of the boys. In 2016 girls' scores were a little higher than the boys', and by 2018 both girls' scores and boys' scores were approximately 7.5 percentage points higher than scores in 2014. In 2018 girls' scores had fallen slightly behind boys' scores again and the gap was back to 3.3 percentage points.

By S2, the questions are more demanding than for lower levels. Many questions use challenging mathematical vocabulary including '*sin*', '*mean*' and '*vector'*, and pupils are required to have knowledge of geometric formulae, for example *Pythagoras' theorem* and the volume of a cube. All questions continue to be multiple choice. Pupils scored more highly on questions involving number operations. Pupils found data interpretation and algebra questions more challenging.

TABLE 9 AVERAGE SCORES IN SECONDARY 2 NUMERACY

Year (S2 Numeracy)	Girls' average score	Boys' average score
2014	27.7%	30.9%
2016	27.5%	26.9%
2018	35.2%	38.5%

3.3 Literacy

3.3.1 Literacy Curriculum

As with numeracy, the literacy assessments were designed to reflect the curriculum in South Sudan, focusing on testing the comprehension of complex passages, rather than reading fluency.

All literacy assessments were based on three texts, with multiple choice question related to:

1. The context of reading:
    a. Reading for literacy experience, including exploring themes, events, characters and settings and bringing the readers' own experiences and knowledge to activities, such as anticipating events and predicting consequences. Texts used could include short stories, poems, plays, biographies or folktales.

b. Reading for information, involving texts such as newspapers, textbooks and magazine articles. Important aspects of such texts are organisation and presentation of information.

2. Aspects of reading:
    a. General understanding shown through knowledge of vocabulary or explaining the purpose of a passage.
    b. Interpreting information, perhaps by comparing two texts or by providing evidence for a specific thought or action.

Examining content and structure, requiring the reader to evaluate the text, to understand humour or to comment on the usefulness of a text for a specific purpose.

3.3.2 Literacy Results

At P5 there was a gap of 5.7 percentage points between boys' and girls' scores at the Baseline, which had decreased to 2.1 percentage points at the Endline. Since the Baseline, scores have dropped by 10.4 percentage points for girls and 13.9 percentage points for boys. At P8, girls' scores were not significantly different from that of the boys in 2014; however this gap widened in 2016 to 2 percentage points but narrowed again in 2018 to 1.7 percentage points. The S2 scores for both girls and boys decreased significantly between 2014 and 2016, but sharply rose in 2018, although not quite to the 2014 level. It is not clear why this is the case, but may reflect the fluctuations potentially inherent in a small sample size. The gap between the average girls' scores dropped from approximately 3% in 2014 and 2016 to 1.9% in 2018.

3.3.2.1 P5 Literacy

P5 literacy assessments involve pupils reading straightforward texts. In P5 all the questions involved reading for information from a first-person narrative, a third-person narrative and a report. There were 16 questions. Where questions were longer or needed some interpretation, pupils tended to score lower.

There was a drop in scores between 2014 and 2016, which continued in 2018. The gap between girls' scores and boys' narrowed from 5.7 percentage points in 2014 to 2.1 percentage points in 2018.

TABLE 10 AVERAGE SCORES IN PRIMARY 5 LITERACY

Year (P5 Literacy)	Girls' average score	Boys' average score
2014	51.6%	57.2%
2016	47.6%	50.3%
2018	41.2%	43.3%

3.3.2.2 P8 Literacy

P8 Literacy Assessments ask pupils to read longer texts. Two narrative texts and one information text are included. Scores were higher for questions relating to the narrative text (70% in 2014, 69% in 2016, and 59% in 2018) than those relating to the to the information text (55% in 2014, 53% in 2016 and 49% in 2018).

In 2014 girls' scores were not significantly different from that of the boys. The gap widened in 2016 to 2 percentage points but narrowed in 2018 to 1.7 percentage points.

TABLE 11 AVERAGE SCORES IN PRIMARY 8 LITERACY

Year (P8 Literacy)	Girls average score	Boys average score
2014	61.5%	62.3%
2016	62.4%	64.4%
2018	53.1%	54.8%

TABLE 12 AVERAGE SCORES BY QUESTION TYPE IN PRIMARY 8 LITERACY

Year (P8 Literacy)	Narrative text average score	Informational text average score
2018	59.1%	48.5%

3.3.2.3 S2 Literacy

S2 literacy assessments involve pupils interpreting more challenging texts.

The scores for both girls and boys decreased between 2014 and 2016, but rose again in 2018, although not quite to the 2014 level. It is not clear why this is the case but may reflect the fluctuations potentially inherent in a small sample size. The gap between the average girls' scores and those of the boys dropped from approximately 3 percentage points in 2014 and 2016 to 1.9 percentage points in 2018.

For S2, the three texts used were an information text about spiders from P8 and two narrative texts, one a local story, and one more challenging, adapted from a novel. For S2, questions are more complex than the primary levels, moving away from retrieving simple information to understanding the way language is used and interpreting the text, i.e. 'Why do you think?' and 'How do we know?' questions.

TABLE 13 AVERAGE SCORES IN SECONDARY 2 LITERACY

Year (S2 Literacy)	Girls' average score	Boys' average score
2014	50.5%	53.5%
2016	30.5%	33.7%
2018	47.5%	49.4%

3.4 GESS Indicators

At the Endline, an increase in the mean test score by 0.25 standard deviations (of the mean of the Baseline test) is a key indicator for the GESS programme. The results are shown in the following table.

TABLE 14 GESS INDICATORS PROGRESS

	Girls				Boys			
	2014 mean score (SD)	**2016 mean score (SD)**	**2018 mean score (SD)**	**Met target? (+.25SD of 2014 mean)**	**2014 mean score (SD)**	**2016 mean score (SD)**	**2018 mean score (SD)**	**Met target (+.25SD of 2014 mean)**
P5 Literacy	51.6% (23.1)	47.6% (20.1)	41.2% (19.5)	NO	57.2% (22.1)	50.3% (19.5)	43.3% (20.6)	NO
P8 Literacy	61.5% (16.8)	62.3% (18.4)	53.1% (18.9)	NO	62.3% (15.4)	64.4% (16.2)	54.8% (17.6)	NO
S2 Literacy	50.5% (13.2)	30.5% (16.7)	47.5% (15.4)	NO	53.5% (11.4)	33.7% (17.9)	49.4% (15.1)	NO
P5 Numeracy	37.5% (15.8)	48.6% (20.0)	41.2% (19.5)	NO	43.5% (17.2)	49.7% (19.0)	41.2% (18.7)	NO
P8 Numeracy	37.1% (15.5)	43.4% (18.5)	35.3% (15.8)	NO	40.2% (15.6)	41.6% (17.6)	35.7% (16.1)	NO
S2 Numeracy	27.7% (11.3)	27.5% (13.5)	35.2% (21.0)	YES	30.9% (12.8)	26.9% (12.2)	38.5% (21.6)	YES

Only in S2 Numeracy did students increase test scores by more than 0.25 standard deviations of the 2014 mean. Girls' scores also increased in P5 Numeracy but not by the required amount.

These results indicate that while GESS has had success in getting girls in to education (enrolment, attendance, etc.), the quality of education received in schools has not significantly changed over the course of the Programme; this is presumably linked to the fact that while GESS interventions tackling enrolment have been rolled out nationwide, those tackling education quality have functioned as pilots, carried out only in a subset of schools.

4. Conclusions

4.1 Contextual considerations of Learning Assessments

The five-year period covered by these Assessments has been a time of great uncertainty in South Sudan, and the Learning Assessments were designed in a different context to that in which they were carried out. For future research, it would be preferable to assess learning beginning from the mastery of basic literacy and numeracy skills.

In 2014 and 2016 fewer pupils and students were enrolled in and attending school, meaning that those pupils were in effect self-selected from families who were economically able to prioritise education for their families; this is particularly the case for girls, whose education is traditionally less likely to be prioritised by their families. In 2018, the larger total enrolment means that newer pupils may be slightly less committed and/or supported by their families than those enrolled in in 2014 and 2016, and this may be a factor in the lower scores seen.

Moreover, in 2018 the date for Assessments was three months earlier than the assessments in 2014 and 2016, giving the pupils three months less education before the test. The difficult economic situation in 2018 may have also caused distractions for teachers and families, with pupils possibly spending less time in school. Continued insecurity in some areas has resulted in disrupted schooling due to displacement, and this may also have contributed to the 2018 fall in scores. However, despite these difficulties, the S2 numeracy scores in 2018 rose above the level of the scores in 2014, and the progress indicator for GESS was met in this particular assessment.

GESS has had a positive impact on narrowing the gap between girls and boys, particularly at P5. Numeracy and literacy results at S2 have increased significantly from 2016; however, consolidated results for all years in both subjects show that performance is down in 2018. A large aspect of this is likely to be due to the assessments being conducted three months earlier in the school year than previously: given that the assessments are curriculum-based, pupils will not have had the same exposure to the curriculum as those in 2014 and 2016.

A key focus for GESS2, MoGEI and education partners should now be to support pupils returning to primary education after either a break of some years and/or with uneven foundational skills. Targeted remedial plans and materials should be put in place to support these students as they join or rejoin education. As in many contexts, a balance needs to be reached between teaching to ensure mastery, and teaching to 'cover the syllabus' – the latter being, in practice, above the current capabilities of many pupils.

4.2 GESS interventions with respect to enrolment appear to have had more impact than ones with respect to quality

Reported enrolment has reached record highs each year that GESS has been operating, reaching over 1.7 million in 2018, from 0.9 million in 2014. Evidence from the GESS Difference in Difference analyses highlights the impact of GESS enrolment- and attendance-focused interventions. By contrast, the Learning Assessment shows less clear evidence of progress in learning: pupil attainment is lower in many of the papers over the course of the Programme. This could be due to a number of reasons, as discussed above. A key point to highlight is that the GESS activities that aimed at improving the quality of education – specifically Teacher Professional Development - were pilots, targeting only 200 schools.

As access to education is broadened and in context of these results, it is vital for future programmes to check that fundamental skills are being learnt at all levels of schooling, rather than starting to assess learning at P5, and focusing on the curriculum. This approach would protect the assessment from being impacted by the time of year it is conducted.

The roll-out of the new South Sudanese curriculum in the near future implies that if learning assessments under GESS2 remain linked to the curriculum, they will need to be redesigned.

4.3 Improving literacy and numeracy results: focusing on interpretation and use of skills

More focus is required on reading and interpreting informational text at all levels of literacy instruction. Classroom materials should expose pupils to a range of different question types when they are expected to interrogate texts. General vocabulary work in the classroom needs to be improved: the use of figurative language is one example where pupils are struggling.

At primary level, increased classroom time for studying the practical applications of mathematics and applying mathematical skills to real-life situations is likely to deliver improved results. At secondary level, more focus is required on interpreting numerical data and identifying mathematical patterns.

These improvements across literacy and numeracy can be led by MoGEI's expertise. For GESS2, Learning Assessments should concentrate on fundamental skills, rather than measuring learning against the curriculum.

Annex

Subsamples of Testing

The following tables show the aggregate test results among only those States and schools that were included in both the 2014 and 2016 samples, i.e. excluding the former States of Jonglei, Unity, and Upper Nile.

TABLE 15 AVERAGE SCORES OF STATES INCLUDED IN ALL 2014, 2016, AND 2018 SAMPLES

	Girls	Boys
2014	44.2%	46.7%
2016	46.7%	45.5%
2018	43.3%	44.6%

TABLE 16 AVERAGE SCORES AT SCHOOLS INCLUDED IN ALL 2014, 2016, AND 2018 SAMPLES

	Girls	Boys
2014	45.2%	47.1%
2016	47.7%	45.0%
2018	42.4%	44.2%

State-level Disaggregation

The Learning Assessment sample was not designed to be representative at the State level. The following data is therefore not sufficiently robust to be indicative of any significant differences between individual States, but has been included as a point of reference.

TABLE 17 AVERAGE TEST SCORES BY STATE

	Girls			Boys		
	2014 (% score)	**2016 (% score)**	**2018 (% score)**	**2014 (% score)**	**2016 (% score)**	**2018 (% score)**
Central Equatoria	48.1%	48.8%	45.1%	51.2%	49.2%	46.1%
Eastern Equatoria	41.3%	36.1%	39.3%	43.5%	42.3%	41.2%
Jonglei	N/A	51.6%	38.8%	N/A	50.8%	48.1%
Lakes	56.5%	52.5%	52.7%	48.7%	49.9%	52.4%
Northern Bahr el Ghazal	36.6%	45.6%	48.8%	39.6%	43.0%	50.2%
Unity	N/A	51.2%	44.6%	N/A	51.4%	46.7%
Upper Nile	N/A	29.4%	33.3%	N/A	29.7%	34.9%
Warrap	41.6%	54.5%	40.6%	46.0%	49.4%	42.9%
Western Bahr el Ghazal	37.2%	31.4%	30.4%	47.9%	31.3%	34.4%
Western Equatoria	33.0%	42.4%	40.2%	39.0%	45.7%	41.1%