Project Evaluation Report

Report title:	Building Girls to Live, Learn, Laugh and 'SCHIP' in Strong, Creative, Holistic, Inclusive, Protective, Quality Education: Midline 1 Report
Evaluator:	VIVA EET Team
GEC Project:	Building Girls to Live, Learn, Laugh and 'SCHIP' in Strong, Creative, Holistic, Inclusive, Protective, Quality Education, VIVA in partnership with CRANE, Uganda
Country	Uganda
GEC window	GEC-Transition
Evaluation point:	Midline
Report date:	August 2020

Notes:

Some annexes listed in the contents page of this document have not been included because of challenges with capturing them as an A4 PDF document or because they are documents intended for programme purposes only. If you would like access to any of these annexes, please enquire about their availability by emailing uk_girls_education_challenge@pwc.com.

Girls' Education Challenge

UKaid from the British people

Evaluation Report for Midline 1 of project 6595 of the Girls' Education Challenge Fund Transition Window

for Grant Holder

Viva

in partnership with

CRANE

(Children at Risk Action Network)

and Fund Manager

Pricewaterhouse Coopers

January 2020

(Revised Aug 2021)

“I know that many of the challenges which hold back progress on learning are highly entrenched in education systems and that achieving change will be difficult. But there is no alternative if we are to bring an end to the learning crisis – so we are up for the challenge.”

Penny Mordaunt, Secretary of State, DFID – DFID Education Policy, 2018.
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/685536/DFID-Education-Policy-2018a.pdf

Project Progress Summary

MAIN POINTS

Overall barriers

Overall the project has done well between Baseline and Midline-1. The Household Survey asks carers if, in the last year, it has become easier for their girl to attend school. Significantly more Intervention carers respond positively compared with Control. The most common reason given is a reduction in the financial barrier; the second most common reason is the enthusiasm of the girls. See 2.2, p14.

Transition Outcome

Transition rates are very high especially for In-School Girls where 97% transitioned successfully between Baseline and Midline-1. Girls who attend a CLC are mostly from poorer families and have had a weaker school experience. The CLCs are effective in promoting successful transitions for 93% of the girls who attend. See 4, p46.

Sustainability Outcome

The Sustainability Scorecard demonstrates significant progress in all indicators. Most progress has been made in levels of engagement between schools, parents, girls and community leaders. The indicators also show important learning on issues of child protection. See 5, p58.

Learning Tests

Project girls significantly outperform Control girls in Learning Tests when the data are analysed in the Benchmark approach and also when the Intervention sample is defined as girls who have been in a project school throughout the time between Baseline and Midline-1. The project is having a greater impact in Literacy than in Numeracy.

When the data are analysed within the Outcomes Spreadsheet or in multiple regression analyses by the FM there are no significant differences between Intervention and Control girls. See 3, p22.

GESI

The project has continued to support the girls recruited through its poverty focus and has performed well in specific work with girls with disabilities. There has been an increase in work with Karamojong girls who face ethnic discrimination. See 2.3, p16

Attendance

Attendance rates are very high wherever reliable data can be obtained. A new phone-based app may help improve the quality of reporting. The main reasons for non-attendance are the financial barrier and illness. See 6.1, p70.

School Management

Improvements are exceeding the targets set in the QIS programme. The scores in the Sustainability Scorecard show positive progress especially around engagement and results from carers in the Household Survey are that management has improved and is very good. See 6.2, p56.

Quality of Teaching

The Household Survey results show significantly smaller proportions of project girls saying that teachers have lower expectations of girls and that they use copying from the board as a teaching method. The Lesson Observation tool results show increasing numbers of teachers satisfying an increasing number of criteria. See 6.3, p83.

Life Skills

The Household Survey results suggest that project girls are more involved in decision-making and appear more confident. Open Qual work shows project girls to have greater levels of ambition and better relationships with their parents and friendship groups. See 6.4, p71.

SUMMARY

Background
The Girls' Education Challenge Transition Project run by Viva in partnership with CRANE works for 9,890 girls who were identified during the GEC 1 project. This GEC Project has been operational since the beginning of GEC (2013) and will continue to the end of GEC-T (2024).

Creative Learning Centres have been a keystone in the work of the project which is based on the identification of girls most likely to miss out on education. The approach to recruitment of GEC girls defines the project and its poverty focus. From the Creative Learning Centres, girls are helped to reintegrate back into mainstream education or in the case of older girls, into vocational training school.

The evaluation consists of a household survey; Learning Tests and qualitative interviews. The EET included some additional qualitative interviews based on the Qualitative Impact Protocol (QUIP™) approach in order to collect observations that were not being obtained from the other tools.

The approach is quasi-experimental with a control population alongside a sample from the 9,890 GEC girls being interviewed at four different evaluation events. Midline is the second of these events, the first to offer comparisons with earlier observations.

Overall project strategy
Tracking individual project girls after the original CLC and school that helped draw them back into the education system may not be sustainable. They are dispersing more than expected and some are now in schools that do not benefit from the work of the project. The work of tracing and identifying them is enormous. An Outcomes Spreadsheet is used to measure the progress in learning between a random sample of girls in treatment schools and control schools. This is a different measure from a sample of girls who have had the benefit of 6 months in a CLC and those who have not. The Outcomes Spreadsheet does not make use of individual difference in difference calculations that might show the true impact of the CLCs. In addition, the Control girls may not be sufficiently similar to the Intervention girls to act as a Control. CLC graduates are not followed sufficiently well to allow learning from this special beneficiary group. A new approach for an evaluation sample should be developed to reduce the workload and deliver more reliable data for measuring the impact of the project interventions.

Learning Outcome Findings
The learning scores are presented in chapter 3.

The results from the Learning Tests were analysed in four different ways, two by the FM and two by the EET –

1. **An Outcomes Spreadsheet** created by the FM provided automatic calculations of changes in average scores for each grade between Baseline and Midline-1;
2. **Multiple regressions** done by the FM to compare the difference-in-difference of Intervention and Control girls while controlling for a range of variables; and
3. **Benchmarking** created by the EET compared the girls' progress between Baseline and Midline-1 against the scores in each grade at the start of the project;
4. **Regressions** done by the EET compared progress made by a "true beneficiary" group in comparison to the Control group. This involved removing from the Intervention sample girls who had not benefited from being in a project school throughout the time between Baseline and Midline-1.

1. The Outcomes Spreadsheet and multiple regressions give insignificant results
The Fund Manager describes below their work on analysing the project data:

After ensuring that the outcomes spreadsheet was populated with data of only those girls who were in the sample at the baseline and at the midline (the successfully tracked sample girls, as per the MEL framework), the FM's quant team was able to successfully replicate scores of the sampled girls, aggregate scores by grade, as well as aggregate scores overall.

The FM team also conducted an unconditional regression analysis, which provided an estimate of the impact for a panel approach – that is, comparing each girl's baseline and midline scores, adding up the differences for each girl, and then calculating how much of the raw score improvement seen between baseline and midline was attributable to the project's interventions.

The EE team was able to calculate the same impact and statistical significance figures as the FM's quant team. The EE team also calculated a range of other regression models, including a benchmarking approach where one grade's progress was compared to that of the grade above them – highlighting some notable successes, such as for primary grade 7. These models are useful for project programming, but don't work for the overall impact figure in the chosen evaluation design, as they do not follow a quasi-experimental panel approach with treatment and control groups and looking at each girls' progress compared to her own scores at baseline.

To increase the accuracy of the impact figure calculated through the unconditional panel regression mentioned above, the FM team asked the EE team to populate tables showing differences in characteristics and barriers between the treatment and the control groups as per the report template. From the report narrative, we know that there are differences between the characteristics of and barriers faced by the treatment girls and those of and faced by the control girls, which are systematic and can influence the improvements in scores seen for each girl. They therefore need to be factored in to the regression analysis in order to isolate project impact as opposed to impact which is a result of the different characteristics and barriers.

Using the data on barriers and characteristics provided by the EE team, the FM identified baseline aggregate scores, baseline grade, Head of Household education level, and the district of intervention as variables which represented characteristics and barriers that are systematically different between treatment and control groups, and which we therefore sought to control through a conditional regression analysis.

The following figures are the output of this conditional regression analysis, which show project impact isolated from any influence that the barriers and characteristics noted have:

Literacy 0.945 SD vs a target of 3.67 SD (26% of target). P-value 0.24. *(The outcomes sheet impact value was 0.81 – see latest outcomes sheet literacy midline tab – but this figure does not provide a panel approach and does not factor in differences between treatment and control groups, which the conditional regression does do.)*
Numeracy -0.116 SD vs a target of 3.01 SD (-4% of target). P-value 0.856. *(The outcomes sheet impact value was 0.65 – see latest outcome sheet numeracy midline tab – but this figure does not provide a panel approach and does not factor in differences between treatment and control groups, which the conditional regression does do.)*

These regressions are based on the data of 587 girls, who were present at baseline and midline, who were out-of-school (but not in a CLC) or in grades P3 to S3 at baseline, and who are either enrolled in school, in a CLC or are still out-of-school. CLC girls are analysed separately, as the panel evaluation approach does not work well in their specific case. *FM Email 28/02/2020*

The FM signed off these figures as the official results of the project.
The results suggest that the project has not had any effect on the performance of girls in literacy or numeracy.

The FM controlled for differences between administrative districts while the EET said in its report that these were confounded by the differences between urban and rural areas. The FM controlled for the main the highest level of education of the PCG using all responses to the question but the EET had found that there were no meaningful differences unless certain subgroups were lumped together.

2. Benchmarking shows significant progress by project girls

The benchmarking exercise compared the progress made by girls in comparison with the average marks in each grade at Baseline. It found that both Intervention and Control girls made more progress than expected and exceeded the levels that the girls in the year above them had at Baseline. However, the differences were very highly significant for the Intervention girls in both Literacy and Numeracy. The differences for the Control girls were not significant.

This suggests that the project girls are making more progress that Control girls and this seems to be particularly true in Literacy. These analyses do not directly compare Intervention scores with Control scores but allow comparisons between the progress made in each group. *The improvements in Learning Test scores are different for different grades but there are bigger improvements in the grades where the project has invested more efforts.*

3. Testing the true beneficiaries shows high impact

The EET discovered that about 200 girls who were considered part of the Intervention sample were not in intervention schools at the time of the Midline surveys. These girls had moved from a project school to a non-project school at some time between Baseline and Midline. The EET removed these girls from the Intervention sample so that it only contained girls who had been in a project school throughout the time between Baseline and Midline and repeated the regression analyses. These showed very highly significant differences between the progress made by Intervention and Control girls. The Intervention girls were making about 2 points more progress than Control girls in a situation where girls had been making on average 4 or 5 points' progress in a year. *The significance of the differences were very highly significant ($p<0.002$). This implies that attending a project school leads to a significant and important improvement in test scores. These results were based on the results of individual girls. The FM believed the results were based on cohort assessments. Therefore the FM declined from using the EET results as the official results. The FM also raised concerns that the date that girl moved from a project school to a non-project school made the alternative results unreliable.*

The girls make progress regularly through the grades but the learning tests show a plateau in marks for girls over the age of 15. The summary in terms of the Outcomes Spreadsheet show that Intervention girls are generally making more progress than Control girls.

The Outcomes Spreadsheet produces results using formulas which are extremely sensitive to small changes in the data. The two tables presented here come from two analyses of almost exactly the same data. The removal of 4 data points changes the p-value from 0.1 to over 0.2.

Literacy		Numeracy	
Beta =	1.69	Beta =	1.19
p-value (1 tail) =	0.134	p-value (1 tail) =	0.184
Target =	2.71	Target =	2.23
Performance =	62%	Performance =	53%

The allocation of the girls' results to grades following two different, but similarly valid approaches, changes the ß values so that the performance changes between 29% and 50 or 60%.

Literacy		Numeracy	
Beta =	0.81	Beta =	0.65
p-value (1 tail) =	0.22[1]	p-value (1 tail) =	0.21
Target =	2.76	Target =	2.26
Performance =	29%	Performance =	29%

The data for the Candidate Year P7[2] where the project invests more resources show the greatest difference-in-difference and significant increases over Control.

Regression analyses of the learning test data tell the same story – in a majority of cases the Intervention girls are making more progress but the results are not significant. The end of Primary grades P5 P6 and P7 show some significant differences in numeracy. It may be possible to argue that a larger number of positive than negative differences can add up to a significant change even where the changes in each individual case are not statistically significant.

School management and teaching quality and girls' safety in school are improving according to the household survey and qualitative interviews. The barriers based on overall lack of resources and negative social norms generally remain in place but are partially lifted for some GEC girls.

Transition Outcome Findings

Better data on girls' grades were obtained by cross-checking variables relating to grades from Baseline and Midline surveys. This produced better assessments of transitions of in-school girls which were above target at about 97%. The numbers are very small (n=7 or 1%) but it seems that GEC girls are more likely to get married or have a baby than Control girls. They are also more likely to continue in school after having a baby. At least 93% of girls who have attended a CLC have successful transitions - most are in school. These high transition rates both for girls in project schools and for CLC graduates represent important successes.

Project heavy investment in P7 teaching seems to have improved transition from Primary to Secondary. The barriers, principally financial, are reduced by work on income generation, savings and work with schools to offer more flexibility in accepting payment of fees.

Sustainability Outcome Findings

The Sustainability Scorecard demonstrates significant progress in all indicators. The indicators on engagement between schools, parents, students and community leaders are

[1] Data on p-values from paired data analyses; other data from cohort analyses.

[2] These are presented under S1 where the girls who spent most time in P7 since Baseline now are.

the most positive. The scoring of Child Protection (three indicators) where work has been intense show small increases largely due to raised levels of awareness and witnesses wishing to revise their assessments at Baseline. The Sustainability Scorecard process needs to be better integrated with other project monitoring work. The QUIP-based qualitative work collected observations of girls with greater confidence and ambition – qualities that support the sustainability of activities and impacts.

Project delivery of transformational change in GESI[3]
Social Inclusion is strong starting with the recruitment of disadvantaged girls into the project – the poverty focus persists through most project activities. Girls with disabilities are supported in mainstream school and through two specialist CLCs – a small number of these girls transition into mainstream school, most get involved in gainful activity. Parents report changed attitudes towards girls with disabilities (GwD) in themselves and their neighbours. There has been increasing support to poor Karamojong girls who also face exclusion based on ethnicity.

The gender approach has been under scrutiny from the fund manager and the project has come through well. A new approach is recommended including work for boys and a review of the practice of focusing entirely on the GEC1 girls. The Gender Framework should be updated and staff expertise consolidated through sharing of the key findings.

Intermediate Outcomes findings
Attendance is shown to be very high wherever reasonably accurate data can be got. The indicators were changed after Baseline and it is difficult to extrapolate from the data collected to the overall beneficiary population or give aggregated results against targets. A new phone-based app shows promise and should be integrated in a new approach to assessing attendance.

School Management improvements are exceeding the targets set in the QIS programme. The scores in the Sustainability Scorecard are very positive especially around engagement and results from carers in the Household Survey are that management has improved and is very good.

The indicator on **Quality of Teaching** based on lesson observations shows progress in line with targets. Girls' and parents' Household Survey responses show significant differences between Intervention and Control on teaching methods and almost all carers say teaching has improved.

Learning Support Teachers are seen to be important in improving the quality of teaching; the level of contract between the school and the household; the monitoring of attendance

Life skills
Intervention girls have significantly higher roles in making decisions than Control girls. They appear to be more confident than Control girls according to their answers in the Household Survey.

The qualitative interviews and the QUIP-based work report very important changes in levels of confidence and ambition among GEC girls and better relationships with their parents, carers and friendship groups. These are important positive findings and will be tested more thoroughly at the next evaluation event.

A new set of questions is required in the Household Survey to improve learning about life skills.

[3] The EET is not required to report under the GESI tool.

Evaluation tools

The Household Survey and the Learning Tests must be overhauled so that they provide better quality observations and quantifiable data for the next evaluation event. More qualitative work should be carried out before the next evaluation.

Outputs

Output indicators – Targets and Midline Status

	Indicator	Midline Target and status
1.1	% girls sampled who made progress in their learning objectives	Target - 70% are making 50% progress
1.2	% guardians sampled who are saving more and report that they are better able to meet school costs	Target 50% Actual - 57%
1.3	% guardians sampled who say they are more involved in their child's education and are now sharing more activities and decisions with their child than before	Target 40% Actual - 759
1.4	% girls sampled who report achieving new competency-based life skills	Target 40% 6354 parents/girls trained 7911 in schools with ICT suites 847 in TVET 99% of 565 girls say ICT lessons good/v good.
2.1	% project-trained teachers who report improving their teaching practice	Target – 20%show all 10 Actual – 53 51% show 5-10
2.2	% girls sampled who describe their school as a safe and inclusive place in which to learn	Target- 30% Actual – 52%
2.3	# schools that improve accessibility for girls with disabilities	Target – 24 Actual – 20 have new accessible toilets
2.4	# schools that improve accessibility for child mothers	Target – 18 CLCs Actual – 10 schools/CLCs run EBCs
2.5	# strengths found by government during school inspections relating to CRANE interventions	Target – at least 2 Actual – many
3.1	% trainees interviewed who describe knowledge and understanding of child safeguarding	Target – 55% Actual – 226 trained
3.2	# communities related to a project school with functional child protection committee that involve parents	Target – 32 Actual – All schools and CLCs received training. All have CP champions. 17 communities have CP committees.
3.3	% trainees sampled who can recite 5 critical steps to take in responding to a safeguarding concern: LISTEN. STOP. REPORT. CARE. RECORD.	Target – 55% Actual – 45 schools/CLCs trained. (=80%)
3.4	Local community leaders meeting together to improve the quality of education and who begin to take over the initiative for leading a collaborative change process. CLC Directors / School heads / DEOs / MoES / Roundtable	Target – Regular meetings Actual – DEOs very term; school directors every year
4.1	# CP policies	Target 52 Actual 45 schools and 4 CLCs
4.2	# Financial policies	Target 52 Actual 48 trained
4.3	# HR Policies	Target 52 Actual 13 confirmed.

Table of Contents

List of Figures

List of Tables

Abbreviations

CCT	Coordinating Centre Tutors
CLC	Creative Learning Centre
CP	Child Protection
CwD	Children with Disabilities
DEO	District Education Officer
EET	External Evaluation Team
EGMA	Early Grades Maths Assessment
EGRA	Early Grades Reading Assessment
FM	Fund Manager
GEC1	Girls' Education Challenge
GEC-T	Girls' Education Challenge Transition
GESI	Gender Equality and Social Inclusion
GwD	Girls with Disabilities
HHS	Household Survey
HoH	Head of Household
IO	Intermediate Outcome
ISG	In-School Girl(s)
IT	Information Technology
JLOS	Justice Law and Order Sector
KCCA	Kampala Capital City Authority
M&E	Monitoring and Evaluation
MoES	Ministry of Education and Sports
OoS	Out of School
ORF	Oral Reading Fluency
OWG	Operational Working Group
PCG	Primary Care Giver
PTA	Parent Teacher Association
PwC	PricewaterhouseCoopers
QIS	Quality Improvement System
qual	Qualitative (interviews or methods)
QUIP ™	Qualitative Impact Protocol
SCHIP	Strong Creative Holistic Inclusive Protective
SeGMA	Senior Grades Maths Assessment
SeGRA	Senior Grades Reading Assessment
SEN	Special Education Needs
SHRP	School Health and Reading Project
TBQ	Transition Benchmark Questionnaire
ToC	Theory of Change
TVET	Technical and Vocational Education and Training
UPE	Universal Primary Education
USAID	United States Agency for International Development
USE	Universal Secondary Education
UWEZO	A project on literacy and numeracy in East Africa
VSLA	Village Savings and Loans Association
WG	Washington Group (definitions of disability)

1 BACKGROUND TO PROJECT

1.1 Project Theory of Change and Beneficiaries

Theory of Change

The project theory of change is based around four main axes which acknowledge the interconnectedness of the many issues that affect girls' education. The axes are defined[4] as:

> *" **Live**: Girls will break the cycle of abuse, violence, exclusion, child marriage, poverty, weak parenting, broken education, and limited literacy. Girls will develop strategies for success and overcoming life barriers through competency-based skills development. Community mentors and engaged parents will support this life journey. Parents and schools will form local clusters that build extra-curricular learning opportunities for children.*
>
> ***Learn**: Girls will achieve enhanced learning outcomes in numeracy, literacy and competency-based skills that orientate them towards purposeful life pathways. Girls will learn new and higher skills inside and outside of the classroom through creative, engaging acquisition of knowledge. Their experience in school will set them on a life pathway of achieving gender equity in the classroom and into adulthood. Girls will be supported by responsible parents, innovative and creative teachers, committed mentors, skilled counsellors and inspiring peers. Their teachers will deliver exciting, quality, inclusive education in child-centred environments that use a variety of methodologies that suit different learning styles, with additional learning support for those who need it.*
>
> ***Laugh**: Girls will overcome the shattering impact of abuse, rejection and failure as psychosocial support and learning therapy builds resilience and confidence. Their schools will do no further harm and help children and families learn how to build safe communities. Local parents' groups will train in holistic wellbeing, parenting, household strengthening, and adult literacy to help provide a smooth transition through to higher education.*
>
> ***SCHIP**: Learning will happen in partner schools where GEC girls have gone to help them become SCHIP schools that provide **S**trong, **C**reative, **H**olistic, **I**nclusive, **P**rotective Quality learning environments with stronger educational and technical foundations that enable accelerated literacy, numeracy and competency-based learning that promote confidence and resilience. Girls will be helped to achieve gender equity in the classroom and into adulthood."*

The invented word SCHIP is an attempt to bring together the different components of project work in supporting girls, teachers, headteachers and others. It does not make it easier to present the project work to others.

The Theory of Change does not require progress to be made in any particular order – improvements in teaching, in child protection or in parental support can occur in any order and reinforce each other. Project staff do not believe that changes in attitude must necessarily occur before changes in behaviour[5].

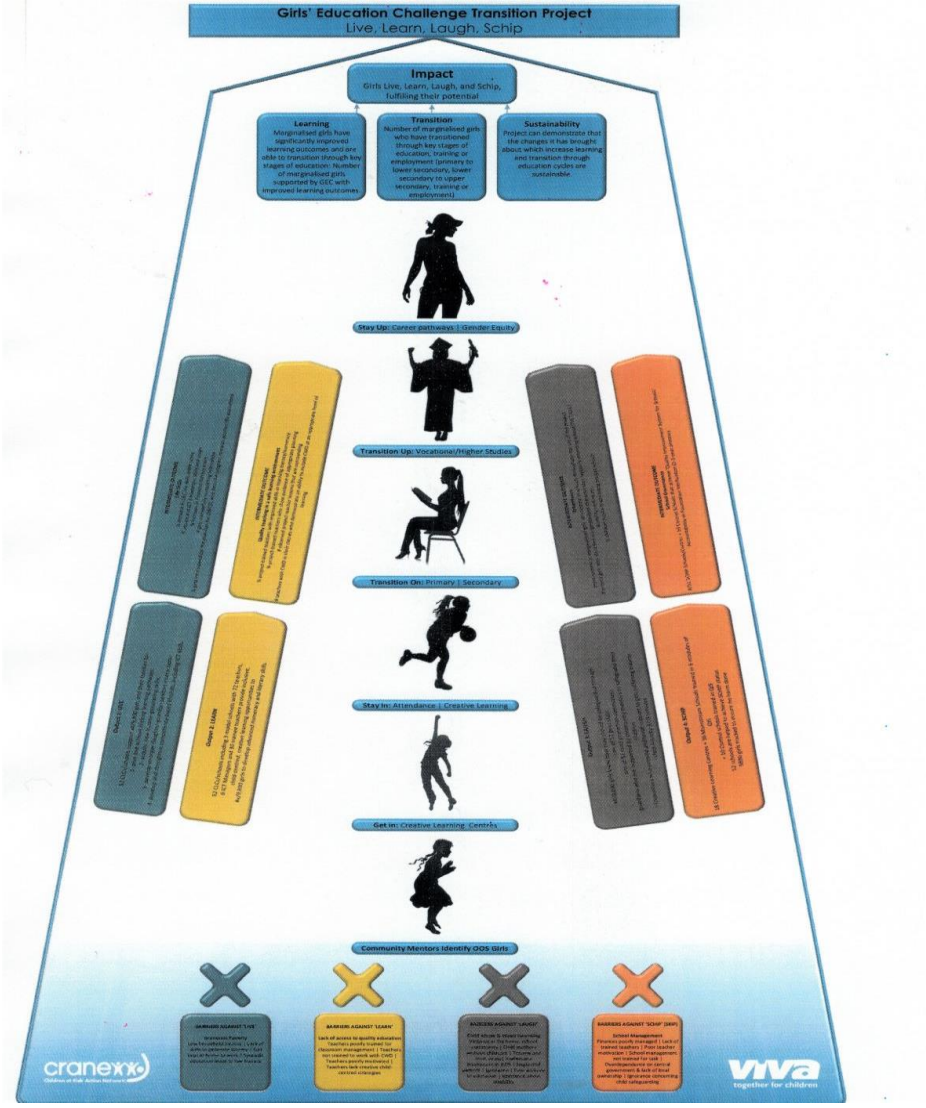
It could be that the project approach is well-adapted to the situation. That is, if girls' education is genuinely complex and it is not possible to predict which intervention will support any particular girl to attend more and perform better in school, then it makes sense to try a wide range of initiatives. For example: one girl may be excited by opportunities in sport; another may go to school to sing or make music; another may decide that school is worthwhile when they encounter positive encouragement from a teacher or do well in a test for the first time; … the list, one could argue, is endless and so the range of interventions should be as wide as possible. The fact that the approach makes monitoring and evaluation more difficult would, in the same argument, be seen as an unfortunate by-product.

[4] Viva and CRANE, Project Proposal, September 2016, p4.

5 The EET observations in the Open Qual work support this understanding.

The project Theory of Change (ToC) is a continuing work-in-progress. It has proved difficult to find a clear and concise way to present the wide-ranging work of the project. There are very large numbers of activities (the Baseline Report lists 45 different areas of activity[6]) and there are significant overlaps between the different areas. The first ToC was criticised in the Baseline Report because it was expressed in a complicated table (based on the logical framework) that made it difficult to link particular changes to particular activities. A new ToC was produced after the Baseline which attempts to breaks away from the logframe design and show how different project activities support girls in different steps and stages (see Fig 1).

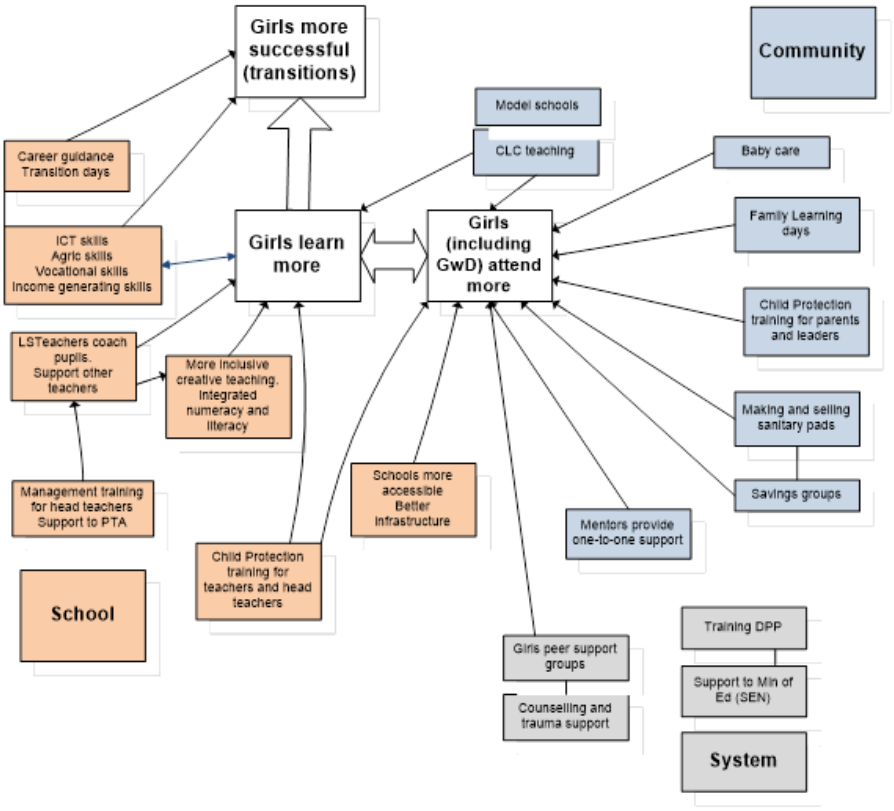
Figure 1 - Project Theory of Change - 2018 scheme



At Baseline, the EET produced a simplified flowchart[7] that showed 19 project activities in three different areas: school, community and education system which contributed to greater attendance or better performance in school which leads to greater life chances. Better attendance and better performance were shown to be mutually reinforcing.

[6] 6595 Viva in partnership with CRANE, Girls' Education Challenge – Transition Window, *Baseline Report (Final) 2018* – Table 1, p8 et seq. We will refer to this document as the Baseline Report.
7 6595 Viva in partnership with CRANE, Girls' Education Challenge – Transition Window, *Baseline Report (Final) 2018* – Fig 4, page 6.

Figure 2 - A version of the Theory of Change from the Evaluation Team



A small version of the flow chart is included here to provide a summary of the approach.

It is hard to imagine a diagram that shows all areas of activity and which provides a clear presentation of the Theory of Change. For the EET, the large number of activities is challenging because it is our job to assess the effectiveness of the project activities even though they do not have results that can be easily separated. At Baseline, the EET recommended a review of activities after the Midline evaluation to try to identify which activities relate to which changes with a view to simplifying the M&E and potentially reducing the numbers of activities. This recommendation has been followed up by the project in discussions with the FM and is repeated in this report in the name of a Post-Midline Review as further simplification would be helpful.

During the writing of this report (September 2019) a new draft ToC was developed by CRANE which is clearer and easier to understand than the previous models. This is largely because it has reduced the number of components so that the entire theory can be seen and read on a single page. The problem of cross-linkages remains so that all eleven outputs contribute to all nine outcomes. The main areas of activity are focused on the family; the school and youth justice. Impact would be assessed in these three areas. The draft is still under development and represents a good attempt to reconcile the need to be comprehensible with the desire to represent all the work of the project. It should be taken forward in the Post-Midline Review.

New activities

Learning Support Teachers had been recruited and some were already deployed to schools where the project was intervening at the time of the Baseline surveys. They had not had any time in post to have had an impact on the project activities at the time of the Baseline. Learning Support Teachers may be a very important addition to the activities. They are shown in this report to be an important force for improving the quality of teaching but also have a role in communication between the school and the household and the monitoring of

attendance. The ability that the LSTs have to pay more particular attention to some students may also be a significant positive contribution to project success. They also represent an initiative that might easily be replicated or extended. They could form an important link between schools and Centre Coordinating Tutors and their cost (one additional post per school) is not beyond the possibilities of local budgets.

Income Generating Activities

More income generating activities (IGA) have been supported since Baseline – some of the increase is just in the numbers involved; some is made up of new activities and some is due to moving some IGA work into schools. IGA work based in schools is common in development programs in Latin America but relatively rare in Africa. There are advantages in that the children are not taken out of school to take part and the activities can be integrated into lessons – finance is mathematics, for example. The increase in IGA work is partly a reaction to the observation that money remains the most important barrier to education. Successful transitions may be only temporary if there is no solution to the financial costs of attending school.

Involving men and boys

The project has extended invitations to men and boys to attend certain discussion groups and there has been some attendance. The group discussions are not new but the topics in some cases are new and aim to draw in the male participants. The question of work with boys has been around since GEC1 and remains unresolved The project is exploring some options of including boys with specific Creative Learning Centres (CLCs).

Violence against children

The publication of the Ministry of Gender report on violence against children (VAC) has added a new impetus to the work of the project with the Office of the Directorate of Public Prosecutions (ODPP). The ODPP analysis is that the worst violence occurs in slum areas around Kampala where populations are more transient.

Inclusion in Karamoja

The project has intensified its activities in Karamoja since Baseline through the work with one of the partners which works with Karamojong street children. The Karamojong are a disadvantaged group that is discriminated against[8] and mostly live in the north-east of Uganda but children move to Kampala and find themselves on the streets. One Mentor has been moved to the area where the children are reintegrated after attending the Creative Learning Centre. There has been more work on Village Savings and Loans Associations, the appointment of Learning Support Teachers and contributions made to improving buildings. This investment is important and forms part of the social inclusion focus of the project. It is not obvious that the current level of investment is cost-effective and the project might assess whether it would be better to invest more (and if so, where to get the additional funds required) or to withdraw from the area.

Beneficiaries

There have been no formal changes in the beneficiary population between Baseline and Midline.

The project beneficiaries are, nominally, the 9,890 girls encountered during the GEC1. The girls fall into three categories: the girls identified by Mentors as needing help in attending or re-joining mainstream school; the sisters of these girls and other girls in the same institutions. The categories are of roughly equal size.

[8] See for example - Cultural Survival (2016) Observations on the State of Indigenous Human Rights in Uganda in Light of the UN Declaration on the Rights of Indigenous Peoples Prepared for the 26th Session of the United Nations Human Rights Council: Universal Periodic Review, March 2016.

Table 1 - Beneficiaries – grades and progressions			
Baseline	Midline	Midline 2	Endline
2018	2019	2020	2022
Primary 1	P2	P3	P5
Primary 4	P5	P6	S1
Primary 5	P6	P7	S2
Primary 7	S1	S2	S4
Senior 1	S2	S3	S5
Senior 2	S3	S4	S6
Senior 3	S4	S5, TVET, work	TVET, work
CLC	CLC, School, TVET, work	CLC, School, TVET, work	CLC, School, TVET, work
TVET	TVET/ work	TVET/ work	TVET/ work
Out of School	CLC, School, TVET, work	CLC, School, TVET, work	CLC, School, TVET, work

Table 2 - Beneficiaries – ages and progressions			
Baseline	Midline	Midline 2	Endline
2018	2019	2020	2022
8	9	10	12
11	12	13	15
12	13	14	16
14	15	16	18
15	16	17	19
16	17	18	20
TVET	17-20	18-20+	20+
OoS	17-20	18-20+	20+

A key subdivision concerns the project girls with disabilities (GwD) who are supported through two Creative Learning Centres (CLC) which provide specialist support. Very few of these children will graduate into mainstream school and their progress is assessed according to their own individual development plans. A number of notable successes have occurred where girls who seemed to have no potential in education have been able to attend mainstream school.

The number of disabled girls being supported by the project has grown since the beginning of the GEC work. It is also true that the GwD in the Household Survey have slightly higher scores for disability at Midline than at Baseline. These changes may be due to increased awareness of parents and teachers having greater skills in identifying disabilities.

The role of the Mentors in identifying the first wave of beneficiaries is pivotal in determining the nature of the project. The Mentors are recruited from the communities local to where the partners are working and are able to engage with the families of girls who they identify as at

risk of failing to make a success of their schooling. This is the starting point for the project and it defines the poverty focus of project activities.

Sub-groups
The EET has worked at identifying sub-groups within the beneficiary population. Most of the characteristics suggested by the FM[9] do not, in themselves, work well in relation to performance in Learning Tests. One issue is that the Mentors' selection methods mean that the households from which the beneficiaries come are among the poorest in each community. This makes the creation of subgroups based on wealth rather difficult. The other difficulty of assessing changes in Learning Test results for different subgroups is that it is necessary to disaggregate results by the age of the girls and by their rural or urban setting. This means that a third characteristic, say, time spent reading at home creates very small subcategories making the results unreliable.

The EET found that a combination of the level of education of the head of the household (HoH) and their main source of income created three subgroups in which performance in Learning Tests looked significantly different and stable. We will use these subgroups in analysing differences between Baseline and Midline.

What is the strategic unit of the project?
The focus on individual girls as beneficiaries has implications for the mechanics of the project and the focus of the evaluation team. Girls change households, locations and schools. These changes make it difficult to follow all the individual GEC girls and it may become inefficient to maintain contact with girls who move significant distances from the main areas of project activity. It becomes difficult to assess the performance of different schools where there are only a few GEC girls or to attribute results to project work in particular communities.

1.2 Project Context

Overall the project context has not changed in significant ways between Baseline and Midline. This is not surprising since the key elements of the context are deep-seated and it has only been eighteen months between the two events.

The Midline Report Template requires this report to stand alone and not be dependent on the Baseline Report. For this reason, we will repeat here the key elements of the project context although they are likely to be familiar to anyone reading this report. There is more detail in the Baseline Report but this section is updated with some findings by the External Evaluation Team (EET) and by more recent reading.

The key elements of the project context are levels of poverty among the beneficiaries; the lack of adequate funding for national educational services and a complex weave of social and cultural norms that make it harder for girls to attend school.

Poverty
The relative poverty of the girls in the project, the "GEC girls", is confirmed in qualitative interviews where wellbeing groups exercises are carried out[10]. Our analyses of the Household Survey also support this categorisation where only about 60 of the 1,100 household heads interviewed at Baseline identify as being in professional employment. The majority identify as farmers with little or no primary education.

[9] For example, in the Midline Report Template.
[10] See Baseline Report - p14 and footnote 29.

General economic indicators suggest that GDP in Uganda has grown since 2016[11] and poverty levels have increased[12]. We have heard of unpublished research in a university in Uganda which shows increasing gaps between rich and poor which seems plausible and makes sense of these overall economic indicators. In any case, the situation of the project beneficiaries has not changed materially.

Low levels of achievement in education

Levels of educational achievement are low in Uganda compared with Tanzania and Kenya. Only 32% of children in P3 to P7 can complete a P2-level literacy and a P2-level numeracy task[13]. Rates are considerably higher in Tanzania and higher again in Kenya. There is no evidence of improvement in the figures for Uganda in the period covering 2011-2015 reported by UWEZO[14].

Missing years of education

One reason for low levels of achievement is that school children in Uganda miss out on years of education as they are withdrawn from mainstream school when their parents cannot or choose not to pay for their attendance. The UWEZO report (2016) shows that the extent of missing years increases with age. Girls who are older or younger than the normal age for each grade perform less well than those who are the age appropriate to their grade. It is also the case that children miss out on schooling because of the absenteeism among teachers – this issue was included in the Household survey.

Differences within Uganda

Children in private schools appear to perform better than government schools according to the UWEZO report. This observation may not be as simple as it seems as it is also true that children in urban schools perform better than those in rural areas. This finding is reported by UWEZO where results are analysed by District[15] and the EET Baseline Report also found that girls in urban locations out-performed rural girls in literacy at all ages[16].

The USAID School Health and Reading Program (SHRP)[17] reports that reading in English and correctly answering questions on passages in English are higher in Luganda-speaking areas than all other areas. Also understanding in English in these areas is higher than understanding in local languages, even where the number of words read is the same. The explanation offered is that people in Luganda-speaking areas "*have higher levels of access to English by way of teachers, print and media[18]."* These findings validate the use of English in Learning Tests in the Viva and CRANE project and may help to explain the higher levels of literacy observed in urban areas compared with rural. It is also important to note how easily different factors can be confounded when assessing levels of literacy.

Education Policy

The policy environment in Uganda is perceived to be positive by most observers. For example, Devries et al, say, "*Uganda has a favourable policy climate, which reflects the government's commitment to address issues of accessibility, participation, capacity-building,*

[11] https://knoema.com/atlas/Uganda/GDP-per-capita
[12] https://knoema.com/atlas/Uganda/Poverty-rate
[13] UWEZO (2017) Are Our Children Learning? Five Stories on the State of Education in Uganda in 2015 and Beyond
[14] UWEZO (2017) Are Our Children Learning? Lessons from UWEZO learning assessments from 2011 to 2015. Dar es Salaam: Twaweza, East Africa.
[15] UWEZO (2016): *Are Our Children Learning?* Uwezo Uganda 6th Learning Assessment Report. Kampala: Twaweza East Africa.
[16] 6595 Viva in partnership with CRANE, Girls' Education Challenge – Transition Window, Baseline Report (Final) 2018 - Tables 49 and 50, page 61.
[17] There is one exception in Runyankore-Rukiga speaking areas for P3 learners mentioned in SHRP, Cluster 3 Follow-Up 3, January 2016.
[18] SHRP, Cluster 3, Follow-Up 3, January 2016, page 4.

awareness raising, care and support of disabled children[19]". Uganda is described in the baseline report of the USAID School Health and Reading Program (SHRP)[20] as having led the way in the promotion of Education for All with the introduction of Universal Primary Education (UPE) in 1997. Universal Secondary Education (USE) was introduced in 2007 at the same time as the Thematic Curriculum. This allowed the teaching in local languages in the first three years of primary school which is seen as progressive and helpful. In 2017, corporal punishment in schools was expressly forbidden in a circular from the Ministry of Education[21] (MoES) citing an amendment to the Children's Act of 2016 and defining corporal punishment as an offence punishable by law. The Circular requires District Education Officers (DEOs) and headteachers to carry out a training program to help eradicate the practice.

Other observers, including some CRANE staff members, express frustration at the failure or slowness with which the government provides practical support for the implementation of its admirable policies. Note that despite the 2017 ban on corporal punishment, half the girls in school interviewed in the Midline survey say that they have seen a child caned at their school in the term in which the survey took place.

Violence against Children

Our understanding of violence against children was informed and updated by the publication in late 2018 of a report from the Ministry of Gender, Labour and Social Development[22]. The survey results in this report are shocking and confirm elements of the project design as appropriate. Rates of reported violence are higher than those found in the Baseline survey interviews. The differences may be due to levels of under-reporting in our interviews.

The Ministry of Gender report led us to other research reports on violence which further improved our knowledge of the context including issues of violence against children with disabilities[23].

These research papers report different patterns of violence that affect boys and girls. Boys suffer more violence at school from peers while girls suffer more violence than boys during their travel to and from school. These observations appear to confirm the appropriateness of lines of enquiry by the evaluation team and activities led by the CRANE project.

Considerable social movement

A feature of the context that has not been stressed before is the extent of social and geographical movement among the beneficiaries. This became evident as the project worked on recontacting families from Baseline with a view to recruiting them for the Midline surveys. Many families had changed location. Some girls had moved to different households often part of the larger family but also sometimes to friends of the larger family.

Is girls' education a complex issue?

Allen *et al* (2016)[24] in their review of assessment methods, describe "*most of the problems*" they encounter as "*wicked problems*" because they are "*complex, multi-faceted, and difficult to solve, and because an effort to solve one aspect of the problem may lead to other, unexpected, difficulties."* This appeal to complexity theory may be important both in terms

[19] Devries *et al*.: Violence against primary school children with disabilities in Uganda: a cross-sectional study. BMC Public Health 2014 14:1017, p8.
[20] USAID/Uganda School Health and Reading Program (2014) The Status of Early Grade Reading and Teaching Reading in Primary School: Cluster 2 Baseline Report, May 2014
[21] Ministry of Education, Circular No. 16/2017, October 31, 2017, Re: Mandatory Response Required - protection of children from all forms of violence including corporal punishment.
[22] Ministry of Gender, Labour and Social Development. *Violence against Children in Uganda: Findings from a National Survey, 2015.* Kampala, Uganda: UNICEF, 2015.
[23] Devries *et al*.: Violence against primary school children with disabilities in Uganda: a cross-sectional study. BMC Public Health 2014 14:1017
[24] Reg Allen, Phil Elks, Rachel Outhred and Pierre Varly (2016) *Uganda's Assessment System: a Road-Map for Enhancing Assessment in Education*, HEART, 14 September 2016

of understanding the project theory of change and in interpreting the results. If girls' education is a complex issue it means that there is no easy link between inputs and outputs and project staff cannot rely on conventional wisdom or "best practice" but need to treat each intervention as an experiment and to react with agility as they learn about what works in specific areas of the context in which they operate.

1.3 Key evaluation questions and role of the Midline Evaluation

The Midline evaluation was designed by the FM to be delivered in 2019. The timing in June was agreed with the FM and fits with the timing of the school terms. This makes the time elapsed since Baseline about 18 months.

The key evaluation questions have been defined by the Fund Manager (FM) and are: do GEC girls make greater or faster progress in literacy and numeracy? and, do they make better transitions at important stages in their schooling and subsequent careers? The Midline Evaluation provides the first opportunity to make assessments of the progress made by the project in these areas.

The EET organised some qualitative work in March 2019 independent of the Midline and Baseline requirements of the Fund Manager. The purpose was to obtain more general impressions of changes that were happening in the lives of the beneficiaries rather than lead interviews with enquiries about the results of the CRANE project work. This work was called Open Qual work to distinguish it from the qualitative work carried out alongside the Midline survey work. The Open Qual results are intended to build on the findings of the Baseline surveys and to guide the work at Midline. The sequence of events (Baseline qual – Baseline HHS – Open Qual – Midline HHS – Midline qual) allows us to build on the learning at each stage.

The project approach has not changed since Baseline. *A longitudinal quasi-experimental approach will be followed. The most important element of the design is the ability to recontact and interview the same girls at the different evaluation stages of the project. Interviewing the same girls at different times allows detailed examinations of changes in their ideas and ambitions and their competencies in learning tests. The approach provides much more robust statistical tests of changes and better understanding of the reasons for observed changes than using group mean scores*[25]*.* Great efforts have been made by project staff to contact the individual girls who took part in the Baseline surveys.

The recontacting tasks were more onerous than expected. The population is more mobile than we had anticipated. Nevertheless, of the 1100 girls interviewed at Baseline, 837 were recontacted and made themselves available for the Midline survey and learning tests. These figures suggest an attrition rate of 24%. See Annex 3 for more discussion of this issue.

The EET has committed a large amount of time in determining the identification of the girls in order to make confident assertions that the same girls are being re-interviewed and re-tested at Midline. This is a very difficult and demanding task as the identity of the girls is made uncertain by variations in names and spelling of names; ages and date of birth; grade or occupation; location; school; family structure and the names and relationships with their primary care giver. Confirming the identity of clients is difficult even where populations are

[25] 6595 Viva in partnership with CRANE, Girls' Education Challenge – Transition Window, Baseline Report (Final) 2018 – page 18.

more stable and it is common to have birth certificates and other identity documents. Even the head of Microsoft Access says it is difficult[26].

The EET has also recommended to the project that significant changes to project activities be postponed until after the Midline Evaluation[27]. This was partly to make the work as consistent as possible during the Baseline to Midline period and to increase the evaluability of the activities. Throughout this period there has been a sense that the project should get over the Midline hurdle before major changes are made.

The Household survey will provide socioeconomic data on the households and information on the ideas and attitudes of the girls and their carers. The most interesting findings will be where there are differences between Baseline and Midline observations. The Learning Tests will deliver information on the competencies of the girls in literacy and numeracy. The project will be judged on whether the overall mean scores of the girls have increased by 0.33 of the Standard Deviation of the Baseline mean scores[28].

The interesting observations will be on where girls have improved their scores and where links can be made between the changes in test scores and the characteristics of the girls, the barriers they face and the project support they have received. That is, the interest will be in whether the learning tests can throw some light on what works in girls' education.

The question of progress in learning tests is straight-forward in its application but complex in interpretation. The assessment of progress is based on mean changes in scores and the mean values are made more robust by creating larger samples and made meaningful by disaggregating into smaller sub-groups. For the project, there are dangers in these processes: analyses that might be meaningful may not be statistically significant and aggregated larger samples may be significantly different but not meaningful.

The qual work delivers information that helps to explain changes that have been noted in the Household Survey results. Ideally the quantitative data will say how things have changed and the qualitative will explain why the change came about.

2 CONTEXT

2.1 Barriers and characteristics

The FM has designed an analysis of Barriers and Characteristics and observations on the interactions of the two different categories. It is not always obvious if an observation like lack of parental support is a characteristic or a barrier but we will try to stick to defining a characteristic as something inherent and the barriers as something external.

The characteristics have not changed significantly since Baseline. We are reluctant to repeat the analyses carried out at Baseline in which eighteen different characteristics were examined for impact on Learning Test results without any being found. The characteristics that do seem to make a difference were those identified as wealth-based subgroups at Baseline. The characteristics of higher levels of education of the PCGs and their attitudes to education and their levels of material wellbeing seem to overlap so that it is not clear if one is

[26] […] *deduplicating/cleaning/scrubbing data that is not in proper relational form is a "decidedly non-trivial task" and one that is unlikely to be easily handled even by a highly skilled database developer/programmer*.' Larry Linson, Microsoft Access MVP. https://bytes.com/topic/access/answers/476564-can-access-use-fuzzy-logic

[27] See Recommendations: 2 (p87), 3 (p88), 4 (p89) and 8 (p90) of the Baseline Report.

[28] See Miske and Joglekar *op.cit.* page 195 for their analysis of the use of 0.02SD as a target for funding decisions in GEC1.

the cause or the effect of another. A good example is reading at home which seems to coincide with higher levels of achievement in learning tests[29]. But it is not clear if this is the case because reading at home occurs more in more wealthy homes where there is more leisure, more light, more encouragement to read and more materials to read. These characteristics may be overlapping because stronger support for education comes from being more educated which may correlate with higher income and so on round and round.

Difficulties (impairments) are characteristics which more consistently map onto educational progress and this is an important issue for those who are affected. However, it is not important for almost all the girls in the project purely in terms of numbers.
Table 3 is mandated by the Fund Manager and attempts to show interactions between Characteristics and Barriers. We have collected the observations from our examinations of Characteristics and Barriers which appear to have an impact on progress in Learning Tests. The Characteristics in the columns relate to the household and the adults in the household. The barriers in the rows relate to the girl or her experience but the distinctions are not always clear. The table shows some of the difficulties with separating cause and effect as mentioned above.

We have done many more analyses which do not produce helpful observations. For example – the girls from households where the head of the household responds to a question about the extent of education that is appropriate for girls by saying that they need only “some secondary” schooling seem to perform less well than girls from households where the HoH responds saying girls should achieve higher levels in education. But the numbers associated with this response are low and when disaggregated across the categories of barriers become meaninglessly tiny.

[29] It may be more accurate to say that not reading at home coincides with lower levels of achievement in learning tests.

Table 3 - (Table 2) Examples of Characteristics and Barriers (numbers of girls/subgroup)																								
	Wealth: Category B				Savings: Not at all				Rural location				HoH education: "No school" or "Some Primary"				Did not pay for school materials this term				Adults alone decide on girl's education			
	C		T		C		T		C		T		C		T		C		T		C		T	
Difficulty hearing	5	7%	14	7%	5	6%	6	4%	10	6%	15	4%	6	8%	30	12%	2	3%	3	2%	8	6%	12	4%
Difficulty remembering	9	12%	21	11%	13	17%	14	9%	23	14%	37	9%	9	12%	31	12%	10	16%	16	12%	13	9%	24	8%
Do not read at home	12	16%	31	16%	6	8%	17	11%	18	11%	52	13%	8	11%	21	8%	4	6%	15	11%	12	9%	19	7%
No one to help	14	18%	32	16%	16	21%	18	12%	35	22%	45	11%	14	18%	28	11%	8	13%	12	9%	22	16%	40	14%
Chores make girl late	16	21%	29	15%	15	19%	22	14%	33	21%	47	12%	15	20%	26	10%	7	11%	12	9%	23	16%	25	9%
Say Most teachers use chalk'n'talk	67	88%	124	62%	62	81%	87	56%	125	78%	215	54%	56	74%	153	59%	44	70%	69	52%	85	60%	159	55%

It is clear that girls in rural areas do less well in Learning Tests than those in Urban areas. One could also call living in the Nakaseke District a Characteristic that makes educational success harder for girls but this is probably because it is mostly rural. In terms of safety however, living in the very urban district of Kampala presents more threats and it could be included as a characteristic except that reporting of lack of safety or fear of lack of safety is relatively uncommon and so it is not possible to link this characteristic to performance or different barriers.

The barrier in the bottom row is about the quality of teaching and refers to the question, "*The main way of teaching is for teachers to write on the board and students copy in their exercise books*." See Table 37 *Barriers – at home and at school*, Section 6.3 and page 80 for more discussion on this question. Where the PCGs agree with this statement the girl tends to score less well in Learning Tests.

The barriers have also not changed significantly since Baseline. The intensity of the barriers may be lessening for some beneficiaries but the fact of the barriers has not changed. Parents still have to find funds to pay for schooling; safety remains a risk; teaching is still often not encouraging and girls still have to do domestic chores. These issues overlap with lower levels of material wellbeing which make it hard to build the sort of analysis that the Midline Report Template suggests. The issues are not separate and cannot with any honesty or to any useful purpose be arranged in a table in which some characteristics interact with some barriers. For example – the characteristic of poverty may be said to interact with the barrier of fees but the two items are just different definitions of each other. The work of the project remains the same.

2.2 Reduced Barriers – more girls love school

The question, "*Over the last year has it become easier for [the girl] to attend school?"* was originally designed to explore changes in access for girls with disabilities but it was asked of all the PCGs whose girl was in school. It yielded a wide range of answers which throw light on changes in barriers.

It also shows a more positive response from Intervention PCGs with 77% saying that things had got easier. The difference between Intervention and Control is significant ($p<0.02$).

Table 4 - Over the last year has it become easier for [girl] to attend school?				
	Yes		No	
Control	148	68%	69	32%
Intervention	385	77%	118	23%

Reducing the financial barrier is the most commonly cited reason for making it easier to attend school for Intervention and Control. It is also the most commonly cited reason for why things have not got easier in both Intervention and Control.

Table 5 - What has made it easier for intervention girls to attend school		
Barrier or change	**Frequency of key words mentioned[30]**	**Overall frequency**
Money and fees	Fees 98, money 15, funds 3, costs 3, dues 3, needs 3, financial 6, loan 1, savings 1, bursary 2, job 1, income 1	159 (41%)
	School materials 4, school requirements 14	
Girls' attitude	Loves school 33, likes school 22, interested 6, committed 3, enjoys school 3, self-motivated 3, more focused 1, always serious with studies 1, concentrates on studies 1, has tried to improve 1, now happily studying 1, always looks forward to school, she understands what she learns,	71 (19%)

[30] We have avoided double counting by cross-checking the statements made against the key words. For example, "pay" is mentioned 35 times but it is not counted when it is in a phrase containing another key word like "fees".

	Given enough care in school, teachers are proud of her, teachers say […] her performance is good, teachers love her too,	
Journey to school	Boarding 36, distance 7, near 14, transport 3, now knows the journey 1, travel with friends, others 2	63 (16%)
Health	Health 12, sick 7, pain 1, itching of the eyes reduced 1, she is ok 1, has received treatment 1, ears clearer 1,	23
Family support	Father supports 6, family support 2, mother does whatever it takes to send her to school 1,counselling from father 1	10
Other forms of support	CRANE (38), help (24), school, headteacher, a teacher, CLCs by name, pastor, a good Samaritan, someone to help me, going to camp 2, an organisation, new friends 2, counselling 2,	

The financial barrier has been reduced in two ways. First, the PCGs have got better access to money and, second, schools have become more flexible in how payments are made. Paying in instalments is specifically mentioned 3 times but the help that is referred to in the last category in Table 5 includes a range of more supportive behaviour and bursars and headteachers being described as more understanding and cooperative.

It is rare that we are able to write the word “love” in a project report but it is the most common word used to explain the second most important change that has made it easier for intervention girls to attend school. In addition to this there are remarks that may help to explain why the girls are feeling more positive in terms of teachers being more understanding and more positive in their relationships with the girls.

The journey to school has become easier apparently by changing school or by moving nearer and in 36 cases by starting to board at school. In one case the role of CRANE in making boarding possible is made explicit.

Improvements in health have also contributed to better attendance but it seems to affect a smaller number than might be expected given that health problems make up 30% of reasons for absence in responses in another part of the Household Survey (See Attendance, Data from the Household Survey).

The pattern for Control families is essentially the same as for intervention but, as explained above, with an overall lower proportion of the population saying that things have improved.

The detailed responses from both Intervention and Control PCGs who said that it has not got easier for their girl to attend school cover the same main points with fees again being the lead reason but this time negatively as paying fees has got harder. The girls are chased away from school in contrast to the situation for those who find it getting better where accommodation is made by the school for more flexible ways of paying.

The journey to school and health issues are also mentioned in that order of priority. Poor health on the part of the parents is also mentioned. There are a few mentions of problems on the way to and from school.

There are virtually no mentions of the girls’ attitudes to school. One PCG says that the girl has become more “unruly” and another says the girl has had a negative attitude. The absence of comments seems to highlight the positive experiences of those girls who are loving school and looking forward to attending.

These observations make clear some of the difficulties in attribution of change to project activities. It's likely that project work on school management and the representation to school managers from Mentors and savings scheme staff have made a difference in how flexible schools have been in allowing PCGs more time to pay school fees. The savings groups are seen as a form of collateral making it easier for schools to be confident that fees will be paid.

It is also reasonable to link the work done with teachers and Learning Support Teachers with the more positive responses of girls to being in school. This would include improvements in classroom teaching and reductions or elimination of corporal punishment. The responses from the PCGs seems to suggest that better teaching methods may be the most important of these changes. It is hard to read the comments from the PCGs without thinking that the whole school environment had become better for girls who have responded by enjoying being in school more than before. This tallies with comments made by Headteachers and Learning Support Teachers with respect to teaching methods and a more friendly environment. It also corroborates the comments made by girls and PCGs during the Open Qual group interviews which revealed more positive approaches to being in school and greater confidence and ambition in the girls.

The project has also probably been instrumental in improving girls' appreciation of school by supporting girls in boarding near or at their school. Boarding is proportionally more important in comments from Intervention areas (10% of comments) than in Control areas (6%).

Similarly, it seems clear that the work by the Mentors in supporting the families and going between the families and the schools has helped to build the levels of support from families and others towards the girls going to school. The links seem more important in Intervention responses than in those from Control partly because there are proportionally more comments and partly because the work of CRANE and the partners is specified in a good number of cases. This reflects the comments made by PCGs in the individual and group interviews during the Open Qual including the development of better relationships and easier exchanges between girls and adults.

2.3 Gender Equality and Social Inclusion (GESI)[31]

The gender dimensions of the context have not changed significantly since baseline but there is an additional focus on the gender approach of the project because of the criticism from the Fund Manager of the EET and project use of the GESI assessment tool. Part of the difficulty has come from the decision that the GEC-T should work only with the girls identified and supported in GEC1. We understood this to be a firm decision and that no changes in the beneficiary population would be allowed.

At the same time the EET had some difficulties with making sense of the GESI assessment tool which was found, despite being called a "continuum", not to be a scale along which project work might progress but a number of different scenarios that project work could be compared with.

We believe that the consultant sent to lead gender training has provided a positive report on the competencies and awareness of the staff. The EET recommended to the consultant that they read the CRANE project Gender Framework[32] which they had not seen.

[31] We have not found an obvious place in the Template for findings on the GESI approach. Our observations are here and conclusions in Chapter 7.

[32] Viva and CRANE, GEC-T Projects Gender Analysis Framework, June 29 2017

During the Open Qual work, the EET interviewed CRANE staff on the gender approach of the project and asked about what gender training they had had. We took part in one day of the training with the visiting consultant and have seen the notes from that day and found in both high levels of awareness and understanding of gender issues in the context in which the project operates and of the effects the project is having on gender relations. Nevertheless, very few staff say they have attended gender training as a separate course and they struggled with the question, *Is working only with girls a gender-sensitive approach*? as if working out the implications for the first time.

It is hard to see how the project can open up new areas of work for boys without increases in funding. The issue should be taken up in a Post-Midline Review.

Social Inclusion started with the recruitment of the GEC1 girls who are among the poorer and more disadvantaged in their communities. This focus remains because of the approach of continuing with the same target beneficiaries. There is also a focus on Girls with Disabilities (GwD) both in the mainstream school system and through two specialist CLCs. The EET pass back to the project the findings on disability from the Household Survey and the more disabled girls who are identified are followed up with additional support. In most cases these girls have already been identified by the Mentors in their communities. The EET has become aware of increased activity in supporting Karamojong girls through work with one partner who specialises in helping street children many of whom come from Karamoja.

2.4 Other barriers

The issues of characteristics and barriers are dealt with in the following chapter in analyses of the effects on the Learning Outcome. We will here examine two barriers made up of questions of safety which do not appear in the following chapter.

The journey to school

The suggestion that girls are more at risk on their journey to and from school is borne out by the responses given by the PCGs in the Household Survey. The figures show over a quarter of responses say that girls are Rarely or Never safe on these journeys. The responses from Control and Intervention are not significantly different.

Table 6 - **Are girls safe on their journey to/from school?**

	Always safe	Usually safe	Rarely safe	Never safe
Control	48%	23%	21%	7%
Intervention	55%	19%	17%	10%

These figures are, however, an improvement on the answers given at Baseline where the Rarely and Never make up 42% of all responses. The difference between Baseline and Midline is very highly significant. The difference remains very highly significant when Intervention and Control observations (which are not significantly different from each other) are combined.

In summary, the situation has changed from one in which carers split their responses:
Baseline – girls are safe on their school journeys 55% Not safe 45%
Midline – girls are safe on their school journeys 74% Not safe 26%

It is not clear why the frequencies have changed. There has been very little action that could directly reduce the risk to girls on their school journeys. The EET has heard stories of teachers taking responsibility for girls' safety where the girl's journey overlaps to some extent with that of the teacher but this is very rare. Again, it may be that the responses really reflect the fact that the carers have thought more about the issue since Baseline.

The situation for boys has not changed significantly although the figures are better at Midline than at Baseline and boys are generally seen to be less at risk on their school journeys. Boys are safe on their school journeys = 80% Not safe = 20%. There are no differences between the responses from Control and Intervention areas either at Baseline or at Midline.

Table 7 - Are girls safe when they go around the area out of the house?				
	Always safe	Usually safe	Rarely safe	Never safe
Control	55%	17%	18%	10%
Intervention	56%	18%	18%	8%

The responses to the question about safety around the home are not different from the question about safety on the way to school.

There are differences between the different geographic areas in which the project works. Locations defined as Urban have lower ratings of safety than rural areas. This is significant (p=0.03).

Table 8 - **Are girls safe around the house? Rural vs Urban**				
	Always safe	Usually safe	Rarely safe	Never safe
Urban	47%	18%	22%	12%
Rural	58%	18%	17%	8%

This difference is also shown in the analyses by District with Kampala being described as the least safe and Nakaseke as the safest. This difference is significant (p=0.01) but other differences are not significant. This is not surprising as Kampala is perceived as busier and more anonymous and with higher rates of alcohol and drug use. We should not immediately trust common sense[33] but the reasons given by the PCGs are consistent, as shown in the following paragraph.

Reasons why journeys may not be safe.

398 respondents chose to give explanations for why girls may not be safe. Most of the causes are because of dangerous and illegal behaviour. See Table 9 below for details.

Table 9 - Why girls may not be safe		
Who poses a danger[34]	Men	86
	Boys	36
	Bad groups, bad company, peers	35
	Boda drivers	11
Crimes or wrongdoing	Kidnap	67
	Rape	48
	Defilement	6
	Sacrifice	5
	Sex	14
	Get girls pregnant	10
Intoxication	Drugs	27

[33] Duncan J. Watts (2011) *Everything Is Obvious, Once You Know the Answer: How Common Sense Fails Us*, New York, NY: Crown. Business. Watts' influential book begins with a story of a social scientist presenting findings to which the audience find obvious explanations. He then reveals that he has lied and that his research showed the opposite of what he has just presented. The audience quickly finds reasons why these results are also obvious.

[34] At Baseline the qual interviews identified danger from bad groups - *The "bad groups" are called by a wide variety of names: "con-men", "distractions", "rolex men", "peer groups", "men calling names", "boys luring us, calling us sexy" and, most common and most negative of all, the motorbike taxi-men, the "boda" riders*. Baseline Report p62.

	Drunkenness	11
Others to blame	Negligent parents	7
	Girls can be tempted	12
	Girls desire things and can be tricked	4
	Road accidents	9
	Accidents not specified	5
	Miscellaneous: “moving at night”, “moving recklessly” “lack of security” “a slum area”	8

Responses to who would you tell?

PCGs were asked who they would tell if they thought a child was being abused either in the community or at school. Responses were not different between Control and Intervention locations with 65-70% saying they would tell the headteacher of the school and a leader of the community. In a community about 20% said they would tell the police. This figure is down from over 30% at Baseline. There is also a small decline between Baseline and Midline in the proportion of those who say they would tell the police about issues of abuse in school.

Box 2: Project’s contribution

The project should respond to the External Evaluator’s comments on the above questions. In particular the project should respond to:

External Evaluator analysis of whether barriers have changed for key subgroups;

The EET has assessed that barriers have not significantly changed, though the impact on individuals has reduced. Key barriers identified are poverty, often expressed as an inability to pay for school fees; safety on the journey to school; teaching styles; the burden of domestic chores.

The Evaluator’s findings that the poverty barrier has reduced is what we would expect to hear from parents who have been involved in savings groups, income generating activities, school enterprises and financial literacy training. We know that not all parents have yet been reached with economic support and the project will not be able to reach all parents because of limited resources. Nonetheless, we will continue to help them to join the savings groups and the school enterprises.

Safeguarding measures in schools and communities have been put in place and at the same time as abuse reducing, the cases of abuse reported is increasing due to a greater awareness and reporting mechanisms now being in place. One Headteacher in a feedback meeting reported to us that now that there are Safeguarding measures in place and the opportunity for whistleblowing, then abuse is reducing because people know that they will not go unnoticed any longer.

Head teachers have reported that teaching pedagogy from the project trained teachers has changed to be more creative and more accessible, and that the children are learning faster now. This is evidenced by the learning test results. We believe that the barrier of poorly trained teachers is reducing and the quality of teaching is increasing. We are aware that the schools still have challenges of genuine inclusion; we have made some inroads into these systemic barriers by getting government approval for the Special Needs Assessment Tool. Over time as this is rolled out across the country, we expect this to help in opening up educational opportunities for children with disabilities and to help identify the needs for adapting learning environments and learning systems.

Access to education for girls with disabilities and social exclusion of children who have been on the streets, and specifically children from Karamoja, have been given specific attention and there is an observable impact on the levels of integration of these girls into the

education system. We believe that the barrier of social exclusion is a real one and will need continued effort and resourcing to overcome.

Many parents have described their increased commitment to getting their children (girls and boys) in school and they are putting a higher priority on education and thus reducing the burden of domestic chores and the timing of when these should be done. The identification of being late to school because of domestic chores and lower learning results is not a surprise, but will give the project leverage to talk to the parents about this specific barrier which should be easy to solve.

We also believe that exclusion from school of girls who for one reason or another have got pregnant is a significant barrier, with most of these girls being expelled from school once they have been found to be pregnant, regardless of the circumstances of that pregnancy. This particular barrier is addressed through the CLCs by allowing pregnant girls and teenage mothers to remain in education whilst their child is cared for as the situation with the girl is worked on with the parents, schools, and if necessary, with the law. This is a unique aspect of the work compared to the control group, with the results below demonstrating that we are tackling this barrier.

Whether contextual changes have an impact on barriers or subgroup;
The impact of poverty as one of the significant barriers to progressive education has been seen impacting more girls in urban areas as families have had to move out of city housing and high-density areas. They have moved to more distant locations where they can afford the rent or have returned to their home villages where they do not need to pay rent. This lightens the financial burden on the family but makes tracking the girls challenging. We have seen many former urban girls move out of the project sphere of influence. This problem is discussed in Annex 17 as something that needs to be addressed by changing aspects of the project design.

Whether activities are still appropriate for subgroups and barriers
We have thoroughly reviewed the subgroups and the barriers to learning with the External Evaluator and have discussed at length the activities which we feel best address the barriers for specific sub groups. We have also considered the most successful activities that contribute to the project's ambitions and have consolidated and revised our project delivery plan.

We are encouraged to see that the results are significantly above 0.33SD in the year groups in which we have put the most effort. We intend to learn from this for what we can do in those critical examination years and see how we can increase impact in other years.

Once again, the model of the CLC has shown that this is a highly successful model for helping to get girls who have been out of school to get back into school. It is a project oxymoron that because no new girls are to be absorbed by the project, one of the most successful aspects of the project which we would wish to demonstrate to other partners are on the verge of phasing out because of their success. We believe that the answer to this is that the CLCs remain open in communities where there are still girls and boys who are out of school or are struggling in school. The learning results show that girls increase in their numeracy and literacy learning rapidly when they are in a CLC and when they remain within the project schools. Therefore the project has entered discussions with stakeholders about how this model can be sustained and shared. In doing this, we need to consider how to share the model of CLCs effectively.

The EET has identified the work with Karamojong children and their specific social needs as worthy of more investment, but that the project needs to decide for itself whether to increase investment or exit the work. Given the impact on the community and the success with the

parents, the girls and with safeguarding in the communities, we are highly likely to give this more attention.

The issue of integration of boys into the model at no significant budget expense remains a barrier to gender equality.

Whether the project plans to review their Theory of Change in light of these findings
We appreciate the comments about our ToC needing to be simplified. We have worked hard with the Evaluator to work out an easier to understand ToC that also makes clearer pathways between activities and outcomes. This is presented in Annex 17. This took some time as we wanted to maintain the output language of the project which staff and beneficiaries have come to understand whilst also making necessary adjustments.

This has reduced the number of activities, partly by grouping similar activities. We are now testing the understanding of the new ToC with staff and stakeholders and will roll it out subsequently.

3 LEARNING OUTCOME

Learning Outcomes in Literacy and Numeracy and mandated by the Fund Manager and the Learning Tests as methods of assessing effectiveness in these areas are defined and approved by the Fund Manager. The project has an Output that states that beneficiary girls will achieve advanced skills in Literacy and Numeracy but the indicators do not include tests in either area of learning. It is unorthodox to have the same aim at Output and Outcome levels but the main point is that the work done at Output level and the indicators of that work cover teaching quality; children's perceptions of safety; access to school for girls with disabilities and demonstrations of approval by school inspectors.

The Learning Tests were maintained exactly as at Baseline in order to make the comparisons between Baseline and Endline as simple as possible. This means that all seven EGRA subtasks and all 8 EGMA subtasks were used. The overall data from Baseline conformed to a Normal Distribution[35] and followed the general rule of delivering higher scores with older girls which made it seem that the tests were adequately effective.

There were nevertheless signs of significant ceiling effects in the Early Grade Tests and the floor effects in the Senior Grade Tests[36]. The same ceiling effects are noticeable in the Midline results and the Learning Tests should be radically overhauled for the Evaluation event 3.

Tables 10 to 17 correspond to the tables 3 to 4b that are designed by the FM in the GEC-T Midline Report Template. The tables contain the numbering from the Template in their title but also have a number which relates to their position in this report.

3.1 Literacy

Mean literacy scores by grade in Table 4 show the gradual increase through the grades which seems appropriate as girls become more competent as they get older. This is the principal observation that supports the use of the literacy tests as a measure of some aspects of the competencies of the girls.

Table 10 - Table 3: Literacy (EGRA/SeGRA)

Baseline Grade cohort	Intervention n	Intervention Group Mean	Control n	Control Group Mean	Intervention Group StdDev
P3	27	20.3	21	18.3	7.4
P4	49	24.5	32	21.4	9.8
P5	57	28.7	29	27.7	12.4
P6	68	34.7	24	32.6	10.5
P7	59	40.6	4	45.9	8.5
S1	30	40.1	27	42.7	13.2
S2	32	47.0	18	45.7	10.0
S3	34	45.6	15	46.5	8.9
S4	32	47.9	0		12.0
S5	5	49.8	0		12.4
S6	4	45.1	0		6.7
OoS	50	36.5	11	33.9	16.6
Total	447	36.2	181	32.5	13.9

[35] Baseline Report, Figure 9.
[36] Baseline Report, Figures 7 and 8.

The Standard Deviations are large compared with the increases by grade. This makes the observation of significant differences between grades or between Intervention and Control means more difficult.

Working on the data makes it obvious how sensitive they are to what seem to be small differences or differences in small numbers of data points. When cleaning the database, we found that relatively minor corrections led to observable changes in mean values in some calculations.

The table also shows a plateau effect in the higher secondary grades. This was noted in Baseline and was interpreted as either evidence that the girls were not making further progress in literacy or numeracy after Secondary 2 or that the tests were not sensitive enough to detect changes. Work on the Midline data showed this phenomenon with smaller and smaller uplifts as the girls progressed through the grades. We decided to change the weighting of the Learning Tests to increase the weight of the Senior Grade tests and reduce the weight of the Early Grade results. We made these changes and recalculated the Baseline Learning Test results with the new weighting. These are the results presented in the tables in this chapter.

Table 5 shows the difficulty mentioned above of finding significant differences between means. First, it is important to note that the girls in intervention and control have made progress between Baseline and Midline. There is only one case where the difference is negative which is the control girls who have moved into Secondary 1. A dip in performance is frequently observed in children in their first year in Secondary which is blamed on the disturbing effects of moving from being the oldest children in the kinder environment of Primary school to being the youngest in the harsher environment of "big" school. This would be a satisfactory explanation except for the result with the Intervention girls who make another increase in performance. This is where we see the largest of all differences between Intervention and Control and there may be a simple explanation as the project invests heavily in girls in Primary 7. There is a great proportion of resources applied because of the importance attached to making the jump to secondary school. This is the point at which girls very often used to drop out. The change in the law creating Universal Secondary Education in Uganda in 2007 has changed the situation but it is still a major barrier especially among the more disadvantaged[37].

Table 11 - Table 3a: Literacy scores from Baseline to Midline

	Intervention				Control				
Baseline Grade cohort	n	Baseline literacy -	Midline 1 literacy -	Difference Baseline to Midline 1 -	n	Baseline literacy -	Midline 1 literacy -	Difference Baseline to Midline 1 -	Difference in Difference (Intervention minus Control)
P3	27	11.5	20.3	8.8	21	8.7	18.3	9.6	-0.7
P4	49	15.5	24.5	9.0	32	12.9	21.4	8.5	0.4
P5	57	18.9	28.7	9.8	29	19.3	27.7	8.4	1.4
P6	68	22.4	34.7	12.2	24	22.1	32.6	10.5	1.7
P7	59	30.8	40.6	9.8	4	42.7	45.9	3.2	6.6
S1	30	32.8	40.1	7.2	27	34.7	42.7	8.0	-0.8
S2	32	37.3	47.0	9.7	18	36.1	45.7	9.7	0.0

[37] https://www.researchgate.net/publication/283641062_Universal_Secondary_Education_USE_in_Uganda_blessing_or_curse_The_impact_of_USE_on_educational_attainment_and_performance

S3	34	37.0	45.6	8.6	15	36.6	46.5	9.9	-1.4
OoS	50	30.5	36.5	6.0	11	29.5	33.9	4.4	1.6
Total	406	25.8	35.0	9.3	181	23.9	32.5	8.7	0.6

The data appear relatively consistent and the girls are apparently making good progress tending to add 8 or 9 points above the year below them. This pattern is though not so consistent among the control girls where there appear some unusual results and it is not clear why that should be.

Table 12 - Table 3b: Literacy results			
Result	**Details**		**Comments**
Literacy Baseline - Midline	Beta =	0.81	Produced using only girls' data who were in both Baseline & Midline1. Pros: paired data, clarity re measurements. Better fit with project approach
	p-value (1 tail) =	0.22	
	Target =	2.76	
	Performance =	29%	P6-7 and P7-S1 stand out for uplift.

Table 12 shows what happens when the data are put into the Outcomes Spreadsheet where invisible formulas carry out calculations that summarise the work into a single figure.

As mentioned in the Summary, the functions carried out within the Outcomes Spreadsheet seem very sensitive to small changes in the data. When five cases were reassigned to their correct Midline grades after being mistakenly given grades above S6 that do not exist the p-value in literacy changed from 0.10 to 0.22. This does not change the EET findings that the p-value is not significant, but it is disturbing and gives us little confidence in the methods being used to assess the project performance.

Similarly, reallocating the girls' learning test marks to different grades by following a different approach halves the assessment of performance from over 60% to less than 30%. The results shown in Table 12 are based on comparing the Learning Test marks of girls based on their grade at Baseline regardless of the grade they are in at Midline. An earlier assessment was based on comparing the mean marks of the girls according to the grade they are in at Midline. If all girls advanced by one grade each year, the assessments should be the same. However, where girls repeat a year their Learning Test marks are likely to depress the mean for the grade they have remained in. See the case of the girls in P6 in the table on page 46 for an example of where this might occur.

The assessments of progress provided by the benchmarking comparisons in the following section avoids these issues by comparing changes in Learning Test marks with the levels achieved in each grade at Baseline.

3.2 Benchmarking Assessment of progress in Literacy

If there were no Control sample, we might try to assess progress against the Benchmark aggregate marks at Baseline. For example – we could expect Girls in P3 to exceed the aggregate mark for girls in P4 by the time they reached that grade if the project work was effective. Exceeding the benchmark set by the grade above would be a measure of success. The benchmark results are shown in Table 13 beside the uplift in marks required to reach that level and the actual level of achievement recorded.

Table 13 - Intervention Literacy Benchmark Assessment				
	Mean of Midline AggRA	Mean of Baseline AggRA	M-B Uplift	B P+1-P Benchmark (Sidelift)
P3	20.3	10.7	9.7	4.4
P4	24.5	15.1	9.4	3.9
P5	28.7	19.0	9.8	3.3
P6	34.7	22.3	12.3	7.5
P7	40.6	29.8	10.8	3.1
S1	40.1	32.9	7.2	4.1
S2	47.0	37.0	10.0	0.7
S3	45.6	37.7	7.9	1.0

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	150.4051	150.4051	42.43982	0.000014
Residual	14	49.61546	3.543962		
Total	15	200.0206			

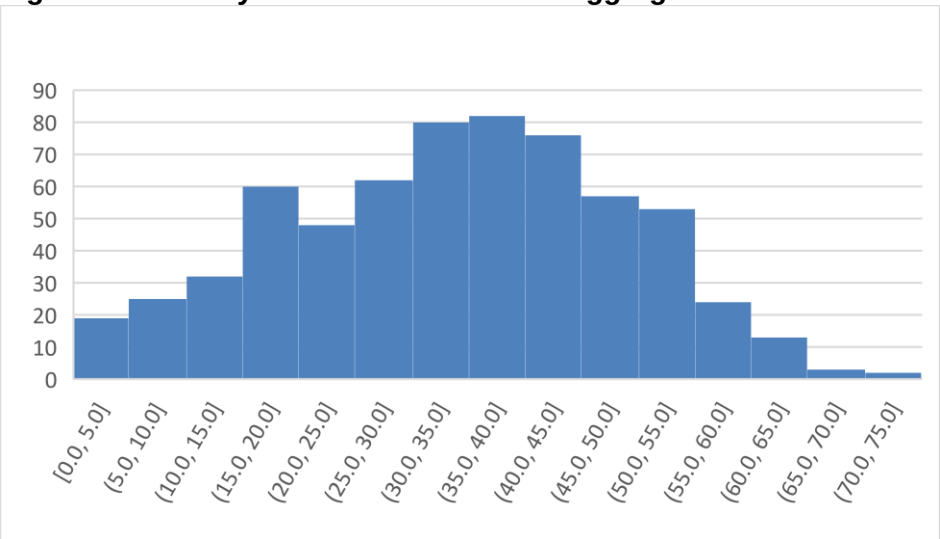
The uplifts in aggregate marks for girls being retested at Midline are larger and sometimes much larger than the differences established in the Benchmark grades at Baseline. The GEC girls could be said to be overtaking their predecessors during the time between Baseline and Midline. The difference is hugely significant.

A similar pattern can be found in the Control sample data – see Table 14. Although the Control girls seem to be making more progress than required to overtake the Benchmark level of the succeeding grade, the differences are not significant.

Table 14- Control Literacy Benchmark Assessment				
	Mean of Midline AggRA	Mean of Baseline AggRA	M-B Uplift	B P+1-P Benchmark (Sidelift)
P3	18.3	8.8	9.5	3.3
P4	21.4	12.0	9.4	6.8
P5	27.7	18.8	8.9	3.6
P6	32.6	22.4	10.2	16.8
P7	45.9	39.2	6.7	-4.7
S1	42.7	34.5	8.3	1.7
S2	45.7	36.2	9.6	1.3
S3	46.5	37.5	9.0	

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	87.25217	87.25217	4.209668	0.060904414
Residual	13	269.446	20.72662		
Total	14	356.6982			

Figure 3 - Literacy overall distribution of aggregate marks



The overall distribution of aggregate marks in literacy resembles a Normal Distribution despite the presence of significant ceiling effects in the Early Grade tests and floor effects in the Senior Grade tests. There may be a slight positive skew but nevertheless it looks overall like an effective spread of results.

3.3 Numeracy

The data from numeracy learning tests tend to have smaller standard deviations than the literacy tests. It is not obvious why that should be the case. Table 15 seems to suggest that SDs get bigger at higher grades. Again, the scores rise sequentially with the grade so it may be that the tests are assessing levels of some numeracy skills.

Table 15 - Table 4: Numeracy (EGMA/SeGMA)

Baseline Grade cohort	Intervention n	Intervention Group Mean	Control n	Control Group Mean	Intervention Group StdDev
P3	27	17.4	21	16.6	5.5
P4	49	21.7	32	19.6	5.8
P5	57	25.7	29	27.5	9.3
P6	68	33.5	24	34.0	10.7
P7	59	39.8	4	34.8	10.3
S1	30	38.6	27	41.3	10.6
S2	32	38.8	18	40.9	8.9
S3	34	45.8	15	43.4	11.4
S4	32	43.9			12.2
S5	5	54.8			12.8
S6	4	57.1			15.0
OoS	50	33.3	11	31.3	12.5
Total	447	33.9	181	30.8	13.3

The table of mean scores by grades with their Standard Deviations invites the question of whether the differences between the grades are significantly different. In many cases, this looks unlikely – for example Control grade S4 is 2.4 points above S3 and both means have Standard Deviations of over 11. Nevertheless, one is tempted to take the increase of 2.4

points at face value because it seems generally true that girls get better at numeracy in the time between S3 and S4.

Table 16 - Table 4a: Numeracy scores from Baseline to Midline									
	Intervention				Control				
Baseline Grade cohort	n	Baseline literacy	Midline 1 literacy	Difference Baseline to Midline 1	n	Baseline literacy	Midline 1 literacy	Difference Baseline to Midline 1	Difference in Difference (Intervention minus Control)
P3	27	11.7	17.4	5.7	21	9.5	16.6	7.1	-1.4
P4	49	15.4	21.7	6.3	32	14.0	19.6	5.6	0.7
P5	57	19.2	25.7	6.5	29	18.1	27.5	9.4	-2.9
P6	68	21.6	33.5	11.9	24	26.1	34.0	7.9	4.0
P7	59	31.0	39.8	8.8	4	33.4	34.8	1.4	7.4
S1	30	35.4	38.6	3.1	27	36.2	41.3	5.1	-2.0
S2	32	34.7	38.8	4.1	18	35.5	40.9	5.4	-1.2
S3	34	39.2	45.8	6.6	15	38.0	43.4	5.4	1.2
OoS	50	29.9	33.3	3.3	11	28.9	31.3	2.4	0.9
Total	406	25.8	32.6	6.8	181	24.5	30.8	6.3	0.6

The table of differences between Midline and Baseline returns the question of whether the differences are significant. The differences between Intervention and Control often look smaller than the differences between grades in the previous table. This seems likely to be the case because improvements in learning brought about by the project are probably not greater than the improvements brought about by another year at school. If the Learning Test results cannot deliver a statistically significant result for girls who have spent over a year in school, is it likely that they can produce statistically significant differences between girls who have spent that year in school receiving teaching of different qualities?

Table 17 - Table 4b: Numeracy results			
Result	Details		Comments
Numeracy Baseline - Midline	Beta =	0.65	0.184 (1-tail). See '[p-Values M1 v02.xlsx]!DiD AggMA' for calc
	p-value (1 tail) =	0.22	
	Target =	2.26	P6-7 and P7-S1 stand out for uplift.
	Performance =	29%	

3.4 Regression analyses of Numeracy data

The regression analyses of numeracy results were carried out following the model provided by the Fund Manager. All the results are provided in Annex XY. The analyses show the Intervention girls doing better than the Control girls in most grades and in P5, P6 and P7 they achieve or approach significance[38].

It is not easy to see what a regression analysis provides in addition to what has been learned from the difference in difference analyses of means. The same story emerges from the Learning Test data: that is, the intervention girls are making more progress in most grades but the differences are slight and data for only a few grades demonstrate a significant

[38] Some are significant at p<0.10 although elsewhere in this report we only recognise p<0.05 as significant.

effect. However, the simple fact that most observations tend in the same direction could imply a significant effect even though the individual tests for each grade do not show a level of significance.

3.5 Benchmarking Assessment of progress in Numeracy

The Baseline results in numeracy can be used as benchmark for assessing changes due to the project work. The EET treated all the benchmarking sample girls in the same way as the learning sample so we have assessments of numeracy skills for all grades at Baseline. The assessment consists of comparing the differences between the mean aggregate marks in different grades at Baseline with the differences in the learning cohort who have made the transition from the earlier grade at Baseline to the new grade in Midline.

The results are shown in table 18. the differences between Baseline and Midline can be seen to be, in general, larger than the differences between successive grades at Baseline.

Table 18 - Intervention Numeracy Benchmark Assessment

	Mean of Midline AggMA	Mean of Baseline AggMA	M-B Uplift	B P+1-P Benchmark (Sidelift)
P3	17.4	11.4	6.0	4.0
P4	21.7	15.4	6.3	3.6
P5	25.7	19.0	6.7	3.0
P6	33.5	22.0	11.5	8.0
P7	39.8	30.0	9.8	4.7
S1	38.6	34.6	3.9	-0.5
S2	38.8	34.2	4.6	5.0
S3	45.8	39.2	6.7	0.4

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	46.53784106	46.53784	6.881339	0.020049314
Residual	14	94.680662	6.762904		
Total	15	141.2185031			

That is the girls have made more progress during their learning between Baseline and Midline than the girls in the year before them had made. In the first row of the table it can be seen that the mean difference between P3 girls at Baseline and the same girls at Midline is 6 points. When the same girls started P3 the girls in P4 were only 4 points ahead of them. Note that the difference is significant.

Table 19 shows the same data for Control girls. Although all the mean marks are increasing the difference is not significant. The Control girls are doing well in terms of the progress they are making but the difference between their progress and the expected progress to come up to the level of the Benchmarks is not as strong as the achievements of the Intervention girls.

Table 19 - Control Numeracy Benchmark Assessment

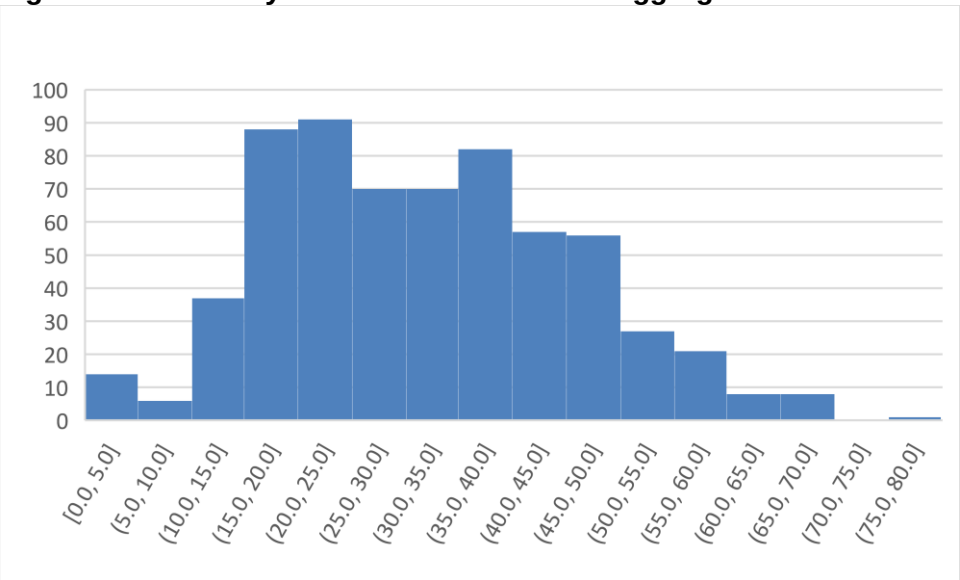
	Mean of Midline AggMA	Mean of Baseline AggMA	M-B Uplift	B P+1-P Benchmark (Sidelift)
P3	16.6	9.3	7.4	4.5
P4	19.6	13.8	5.8	3.5
P5	27.5	17.3	10.2	7.7

P6	34.0	25.0	9.0	4.1
P7	34.8	29.0	5.8	7.2
S1	41.3	36.2	5.1	-1.3
S2	40.9	34.9	6.0	4.9
S3	43.4	39.8	3.6	

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	18.71444666	18.71445	2.878842	0.113548095
Residual	13	84.50891105	6.500685		
Total	14	103.2233577			

This may be, alongside the same result in Literacy, the most convincing evidence of impact of the project work. The EET would make more of the observations if we were more confident in the robustness of the data.

Figure 4 - Numeracy - overall distribution of aggregate marks



The overall distribution of numeracy aggregate marks for girls in school grades P4 to S4 shows an approximation to a Normal Distribution. The actual shape of the histogram can be changed by modifying the bin size and larger bins make a smoother shape but it still looks a slight positively skewed distribution. This may be further evidence to support increasing the weight of the Senior Grade tests.

Skills gaps and grade achieved

The data from the Learning Test subtasks are reworked into two composite indicators that should tell us more about progress being made by girls in numeracy and literacy. The first of these is called Foundational skills gaps. The percentages of girls' scores in the different subtasks are put into categories of different levels of learner. In a normal pattern, the proportions would gradually increase in the lower part of the table as the learners moved from weaker to stronger learners. This is exactly what can be seen in the following tables. The girls are becoming better at numeracy and literacy especially among lower level learners and it may be that it is harder to make progress in the higher learner categories. This observation tallies with the comment in the preceding paragraph on the possibility of the overall distribution being positively skewed and with the observation of a plateau in Learning Test scores among older girls. There are several possible explanations for these observations. First, the girls do not increase their skills in literacy and numeracy above a

certain age and, second, the Learning Tests are not capable of detecting increases above a certain level. The latter explanation may be due the selection of all subtasks used in the Baseline and Midline surveys and the weighting of the Early Grade and Senior Grade subtasks. We adjusted the weighting towards the senior grade tasks which reduced but did not eliminate the plateau effect. These issues will be taken into consideration when redesigning the Learning Tests for the next evaluation event.

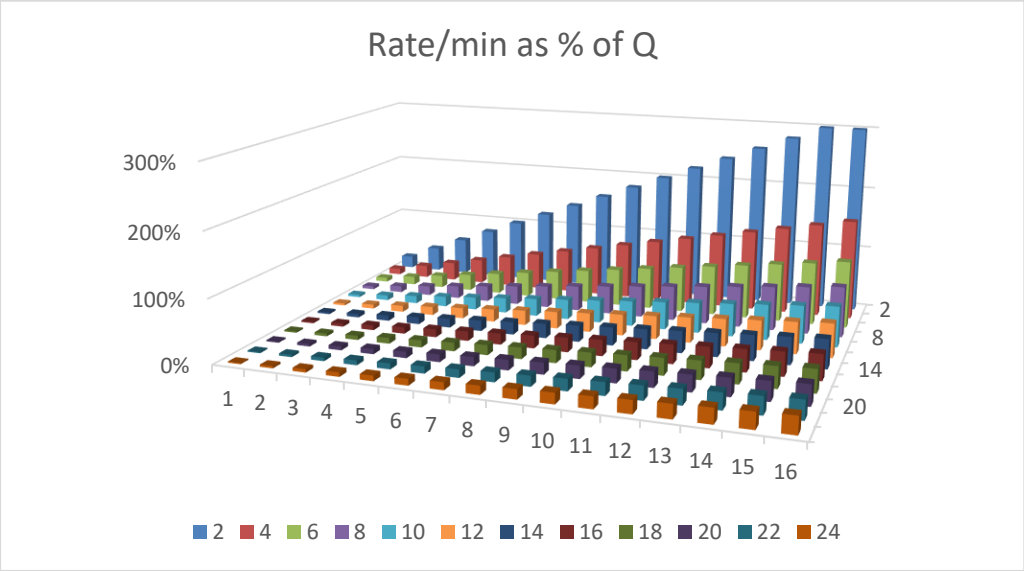
Table 20 - Table 6: Foundational literacy skills gaps											
Categories		M1 EGRA 1 Letter Name Knowledge. Learning Cat (based on Rate)	M1 EGRA 2 Initial Sound Identification. Learning Cat (based on Score)	M1 EGRA 3 Letter Sound Identification. Learning Cat (based on Rate)	M1 EGRA 4 Familiar Word. Learning Cat (based on Rate)	M1 EGRA 5 Invented Word. Learning Cat (based on Rate)	M1 EGRA 6a Oral Reading Fluency. Learning Cat (based on Rate)	M1 EGRA 6b Comprehension. Learning Cat (based on Score)	M1 SeGRA 1 Comprehension + Analysis. Learning Cat (based on Score)	M1 SeGRA 2 Comprehension + Inference. Learning Cat (based on Score)	M1 SeGRA 3 Short Essay. Learning Cat (based on Score)
Non-learner	(0%)	1% (- 0%)	18% (- 2%)	23% (- 15%)	2% (- 2%)	10% (- 10%)	5% (- 8%)	18% (- 1%)	13% (- 34%)	13% (- 40%)	46% (- 24%)
Emergent learner	(1-40%)	9% (- 23%)	14% (+ 3%)	64% (+ 13%)	46% (- 16%)	64% (- 3%)	12% (- 8%)	28% (+ 7%)	41% (+ 6%)	67% (+ 29%)	43% (+ 20%)
Established learner	(41-80%)	61% (+ 4%)	37% (+ 3%)	13% (+ 3%)	50% (+ 18%)	26% (+ 14%)	46% (+ 8%)	47% (- 1%)	45% (+ 28%)	20% (+ 12%)	11% (+ 4%)
Proficient learner	(81-100%)	29% (+ 19%)	32% (- 3%)	0% (- 1%)	2% (+ 1%)	0% (- 0%)	36% (+ 8%)	7% (- 6%)	2% (+ 0%)	0% (- 0%)	1% (- 0%)
		100%	101%	100%	100%	100%	99%	100%	101%	100%	101%

Table 21 - Table 7: Foundational numeracy skills gaps											
Categories		M1 EGMA 1 Number Identify. Learning Cat (based on Rate)	M1 EGMA 2 Quantity Discrimination. Learning Cat (based on Score)	M1 EGMA 3 Missing Numbers. Learning Cat (based on Score)	M1 EGMA 4 Addition. Learning Cat (based on Rate)	M1 EGMA 5 Subtraction. Learning Cat (based on Rate)	M1 EGMA 6 Word Problems. Learning Cat (based on Score)	M1 EGMA 7 Multiplication. Learning Cat (based on Rate)	M1 EGMA 8 Division. Learning Cat (based on Rate)	M1 SeGMA 1. Learning Cat (based on Score)	M1 SeGMA 2. Learning Cat (based on Score)	M1 SeGMA 3. Learning Cat (based on Score)
Non-learner	(0%)	1% (+ 0%)	1% (+ 0%)	1% (- 1%)	2% (+ 0%)	4% (+ 1%)	4% (- 3%)	11% (- 8%)	9% (- 12%)	3% (- 15%)	22% (- 18%)	53% (+ 3%)
Emergent learner	(1-40%)	62% (- 22%)	1% (- 1%)	5% (- 20%)	95% (+ 4%)	89% (- 5%)	17% (- 7%)	88% (+ 9%)	91% (+ 12%)	40% (- 19%)	52% (+ 14%)	39% (- 9%)
Established learner	(41-80%)	37% (+ 21%)	12% (- 6%)	65% (+ 10%)	3% (- 3%)	8% (+ 3%)	35% (- 2%)	0% (- 0%)	0% (- 0%)	54% (+ 31%)	24% (+ 3%)	7% (+ 5%)
Proficient learner	(81-100%)	0% (+ 0%)	87% (+ 6%)	30% (+ 10%)	0% (- 1%)	0% (+ 0%)	45% (+ 12%)	0% (+ 0%)	0% (- 0%)	4% (+ 3%)	2% (+ 1%)	0% (+ 0%)
		100%	101%	101%	100%	101%	101%	99%	100%	101%	100%	99%

The Foundational Skills gaps tables are based on calculations of the rates at which timed tasks are done. This has an enormous effect on the results. The raw data in the simple scores would show very significant ceiling effects in almost all the Early Grade subtasks. The use of rates spreads out the data more and allows a clearer observation of changes over time.

But it is highly doubtful that using timed tests and calculating rates is the best way of assessing girls' competencies. First, our main interest is in the girls' levels of competence rather than their speed. Speed is partly a measure of confidence which isn't really part of the test. If we really wanted to know how quickly girls can do Learning Test subtasks, we should introduce them as speed tests. A highly literate girl might read accurately and slowly because it is more normal and makes it easier to see how correctly she is performing. Second, the speed distorts the rate when girls complete the task very quickly as can be seen in Figure 3. The task in question contains 16 elements and many girls get them all correct. The time taken is shown on the z-axis (depth) and as the time gets shorter the calculated rate gets disproportionately larger. This means that a girl who completes the 16 elements in under 10 seconds but making several mistakes would get a higher rate than a girl who completed all the elements successfully.

Figure 5 - How speed affects Rates calculations in EGMA Subtask 1



The important point here is that the underlying mathematical functions that determine the Skills Gaps and the Grade Achieved results are very sensitive and the results could be changed by minor modifications to the formulas being used. This might not matter if the results are being used to observe changes in performance and provide useful learning to the project but it does matter if the results are used to judge the project and make decisions about its future.

The problems of using timed tests are described forcefully by Miske and Joglekar (2018)[39]. They present evidence that the 60 second time limit is particularly unhelpful in the Oral Reading Fluency subtask. They quote Bartlett et al[40], "*Limiting a reader to one minute does*

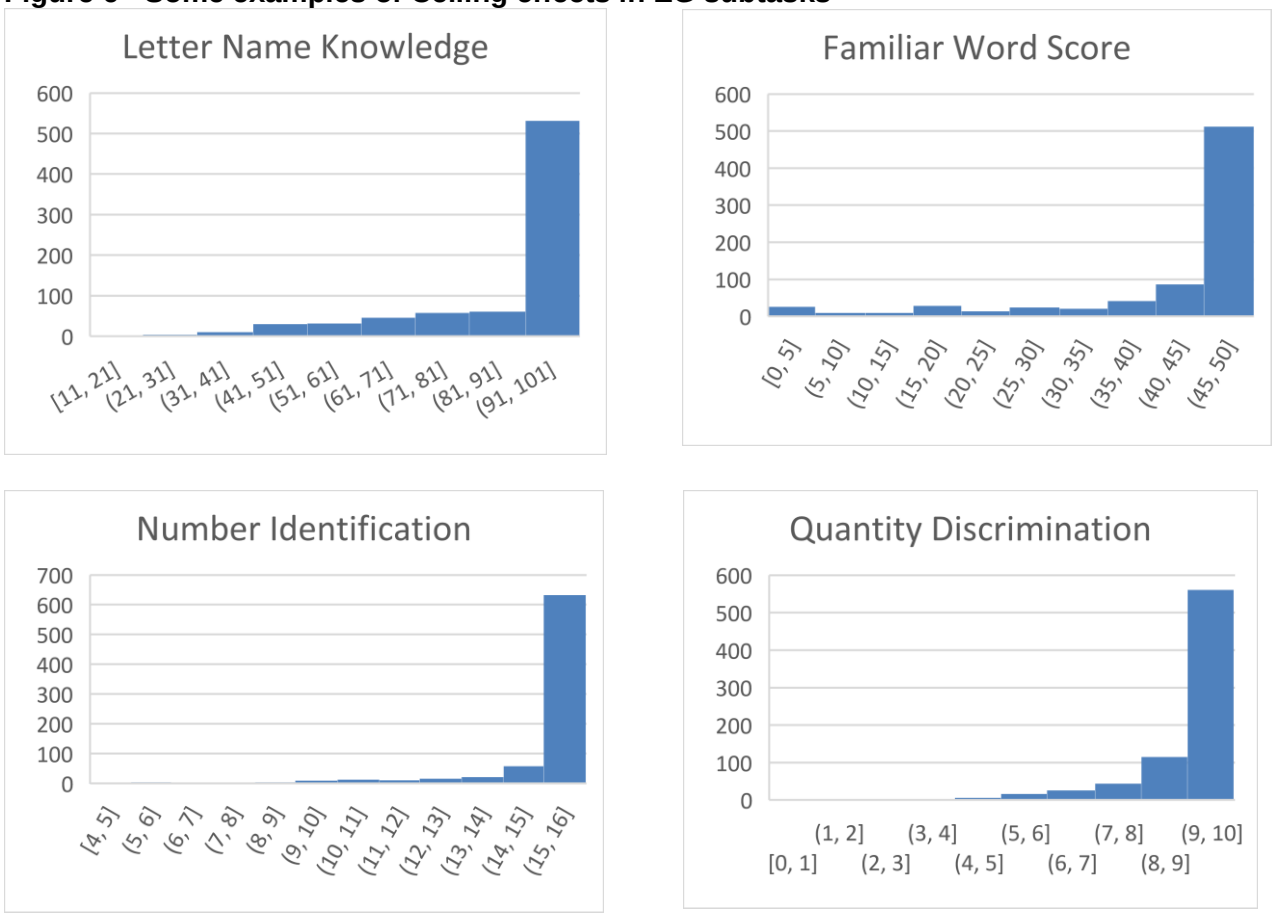
[39] Shirley J. Miske, Alison B. Joglekar, "*Using or Misusing the Early Grade Reading Assessment? Examining A Measure of Payment by Results in the Girls' Education Challenge*" In Annual Review of Comparative and International Education 2017. Published online: 22 Jan 2018; 187-202. p192

[40] Bartlett, L., Dowd, A. J., & Jonason, C. (2015). *Problematizing early grade reading: Should the post-2015 agenda treasure what is measured?* International Journal of educational Development, 40(C), 308–315.

not allow slow readers sufficient time to demonstrate fully their ability to comprehend a passage". They point out that reading aloud without comprehension is not reading. They conclude that the EGRA subtask is a "*highly contested measure of learning to read with comprehension, especially in multilingual contexts*."

The other key point is that the rates methods we have used hide the ceiling effects that are occurring in the Early Grade Learning Tests. Some of the ceiling effects are quite severe and a complete overhaul of the Learning Tests will be necessary before they are used again.

Figure 6 - Some examples of Ceiling effects in EG subtasks[41]



Among the EGRA subtasks Letter Sound Knowledge and Oral Reading Fluency stand out as not having a significant ceiling effect. Initial Sound Identification has only a small ceiling effect. The EGMA subtasks without a strong ceiling effect are Multiplication and Division and Word Problems has only a slight ceiling effect although it only contains 6 elements.

It should be possible to examine the results from all the subtasks and arrive at recommendations for a set of Learning Tests which would allow a continuation of assessments of literacy and numeracy levels which would be much more efficient and involve less ethical jeopardy.

The Grade Achieved calculations are like the Foundational Skills Gaps volatile and dependent on the definitions of the functions used in the calculations. In an ideal

[41] The Y-axis in these graphs show the numbers of girls in each category of points scored shown on the X-axis. In the first example, 532 girls scored above 90 out of 100 in the Letter Name subtask. We have not found a definitive definition of "ceiling effect" but these examples surely qualify.

assessment the numbers of girls would gradually move down the table as the girls get older and further to the right as they became more competent in the Learning Tests. We have included Baseline and Midline results and an additional table of the differences[42] so that this pattern can be observed. It is not easy to see and there appears to be a lot of meaningless movement of the numbers which makes the pattern not very distinct. This is probably due to the initial conditions used to set the Grades Achieved and there may be other definitions that would show the gradual increases in the scores to be seen more clearly.

Grade Achieved – Literacy

Data in the Grade Achieved tables (Table 22 to Table 27) are restricted to the girls who were interviewed and did Learning Tests at both Baseline and Midline.

Table 22 - Baseline Grade Achieved - Literacy

	Sum of Grade 1 Achieved -	Sum of Grade 2 Achieved -	Sum of Grade 3 Achieved -	Sum of Grade 4 Achieved -	Sum of Grade 5 Achieved -	Sum of Grade 6 Achieved -	Sum of Grade 7 Achieved -	Sum of Grade 8 Achieved -	Sum of Grade 9 Achieved -
Primary 3	0	4	1	0	0	0	0	0	0
Primary 4	2	15	1	1	0	1	0	0	0
Primary 5	4	38	4	5	0	0	0	1	0
Primary 6	1	59	10	13	0	9	0	0	0
Primary 7	6	63	11	16	2	6	0	1	0
Senior 1	12	56	11	13	1	3	0	4	0
Senior 2	10	47	9	19	0	8	0	6	0
Senior 3	9	47	10	14	3	12	2	10	1
Senior 4	17	49	14	18	0	9	1	10	1
Senior 5	1	5	2	3	0	0	0	2	1
Senior 6	2	7	2	1	1	0	0	5	2

Table 23 - Midline Grade Achieved - Literacy

	Sum of Grade 1 Achieved -	Sum of Grade 2 Achieved -	Sum of Grade 3 Achieved -	Sum of Grade 4 Achieved -	Sum of Grade 5 Achieved -	Sum of Grade 6 Achieved -	Sum of Grade 7 Achieved -	Sum of Grade 8 Achieved -	Sum of Grade 9 Achieved -
Primary 3	0	0	0	0	0	0	0	0	0
Primary 4	3	23	0	2	0	1	0	0	0
Primary 5	2	43	3	8	0	1	0	0	0
Primary 6	14	79	5	15	0	11	0	1	0
Primary 7	17	68	2	38	0	12	0	3	0
Senior 1	17	56	5	33	0	12	0	4	0
Senior 2	22	54	10	37	2	14	0	11	1
Senior 3	20	47	8	40	0	23	0	13	0
Senior 4	23	53	4	41	4	18	0	15	2

[42] Negative differences indicate Grades Achieved with containing fewer girls at Midline than at Baseline and are shown in red.

Senior 5	8	10	1	9	0	5	0	7	1
Senior 6	2	7	2	6	1	5	0	2	1

Table 24 - Grade Achieved – Literacy changes between Baseline and Midline

	Sum of Grade 1 Achieved -	Sum of Grade 2 Achieved -	Sum of Grade 3 Achieved -	Sum of Grade 4 Achieved -	Sum of Grade 5 Achieved -	Sum of Grade 6 Achieved -	Sum of Grade 7 Achieved -	Sum of Grade 8 Achieved -	Sum of Grade 9 Achieved -
Primary 3	0	-4	-1	0	0	0	0	0	0
Primary 4	1	8	-1	1	0	0	0	0	0
Primary 5	-2	5	-1	3	0	1	0	-1	0
Primary 6	13	20	-5	2	0	2	0	1	0
Primary 7	11	5	-9	22	-2	6	0	2	0
Senior 1	5	0	-6	20	-1	9	0	0	0
Senior 2	12	7	1	18	2	6	0	5	1
Senior 3	11	0	-2	26	-3	11	-2	3	-1
Senior 4	6	4	-10	23	4	9	-1	5	1
Senior 5	7	5	-1	6	0	5	0	5	0
Senior 6	0	0	0	5	0	5	0	-3	-1

Table 24 shows the differences between grade achieved at Baseline and Midline. There seem to be two movements – first a larger movement from Grades Achieved 2 and 3 into Grade Achieved 4. There is a smaller movement from Grade Achieved 5 into Grade Achieved 6. The pattern of a general movement down the table towards the right is discernible although there are numbers in the table that do not fit the pattern.

Grade achieved - Numeracy

Table 25 - Grade Achieved Numeracy - Baseline

Grade	Sum of Grade 1 Achieved -	Sum of Grade 2 Achieved -	Sum of Grade 3 Achieved -	Sum of Grade 4 Achieved -	Sum of Grade 5 Achieved -	Sum of Grade 6 Achieved -	Sum of Grade 7 Achieved -	Sum of Grade 8 Achieved -	Sum of Grade 9 Achieved -
Primary 3	0	1	0	0	0	0	0	0	0
Primary 4	4	0	0	0	0	0	0	0	0
Primary 5	2	0	0	1	0	0	0	0	0
Primary 6	4	1	0	11	0	4	0	0	0
Primary 7	8	2	2	18	0	14	1	1	0
Senior 1	12	3	4	26	3	25	0	1	0
Senior 2	4	0	1	25	0	23	0	1	0
Senior 3	19	4	4	25	2	28	1	3	0
Senior 4	18	6	6	20	1	32	2	4	0
Senior 5	4	0	0	3	0	3	1	1	0
Senior 6	3	1	1	5	3	5	2	2	0

Table 26 - Grade Achieved Numeracy - Midline									
	Sum of M1 Grade 1 Achieved -	Sum of M1 Grade 2 Achieved -	Sum of M1 Grade 3 Achieved -	Sum of M1 Grade 4 Achieved -	Sum of M1 Grade 5 Achieved -	Sum of M1 Grade 6 Achieved -	Sum of M1 Grade 7 Achieved -	Sum of M1 Grade 8 Achieved -	Sum of M1 Grade 9 Achieved -
Primary 3	0	0	0	0	0	0	0	0	0
Primary 4	2	0	0	2	0	0	0	0	0
Primary 5	8	0	0	8	0	0	0	0	0
Primary 6	24	0	0	35	0	6	0	3	0
Primary 7	30	1	3	51	2	19	0	10	0
Senior 1	24	0	5	49	2	24	1	7	0
Senior 2	25	0	5	46	2	26	1	6	0
Senior 3	21	0	2	37	1	21	1	1	0
Senior 4	32	4	6	50	7	30	4	12	1
Senior 5	10	1	2	10	3	9	1	3	0
Senior 6	7	0	3	7	1	7	1	2	0

Table 27 - Grade Achieved Numeracy – Changes between Baseline and Midline									
	Sum of M1 Grade 1 Achieved -	Sum of M1 Grade 2 Achieved -	Sum of M1 Grade 3 Achieved -	Sum of M1 Grade 4 Achieved -	Sum of M1 Grade 5 Achieved -	Sum of M1 Grade 6 Achieved -	Sum of M1 Grade 7 Achieved -	Sum of M1 Grade 8 Achieved -	Sum of M1 Grade 9 Achieved -
Primary 3	1	2	3	4	5	6	7	8	9
Primary 4	0	-1	0	0	0	0	0	0	0
Primary 5	-2	0	0	2	0	0	0	0	0
Primary 6	6	0	0	7	0	0	0	0	0
Primary 7	20	-1	0	24	0	2	0	3	0
Senior 1	22	-1	1	33	2	5	-1	9	0
Senior 2	12	-3	1	23	-1	-1	1	6	0
Senior 3	21	0	4	21	2	3	1	5	0
Senior 4	2	-4	-2	12	-1	-7	0	-2	0
Senior 5	14	-2	0	30	6	-2	2	8	1
Senior 6	6	1	2	7	3	6	0	2	0

Patterns in changes since Baseline are harder to see in Table 27 of changes in Grade Achieved in Numeracy. There seems to be a general movement from Grade Achieved 1 to Grade Achieved 4 but it is not clear how all the scores have changed. There is also a movement from Grade Achieved 6 to Grade Achieved 8. As we have pointed out the data seem sensitive to small changes and it may be that this analysis is asking a lot of some data that are not robust enough to deliver.

3.6 Subgroups by characteristics

The survey sample is spread across four different districts in the Central Region of Uganda. The Learning Test means show some variations in literacy and smaller differences in numeracy. The results are directly in line with the findings of the UWEZO studies[43].

Table 28 - Learning test results by location				
	Mean Literacy Mark	Change in Literacy mark	Mean Numeracy Mark	Change in Numeracy Mark
a) Kampala	41.6	15.9	36.8	12.3
b) Wakiso	42.7	10.8	38.3	7.1
c) Mukono	37.5	12.9	35.1	9.9
d) Nakaseke	37.1	14.3	35.5	9.5
Rural	38.2	12.1	35.7	8.5
Urban	41.3	17.0	36.8	13.8

The learning test results at Baseline showed a generally higher level among girls from Urban areas compared with rural. This observation was also made by the USAID School Health and Reading Program[44]. The differences seem to continue into the Midline survey and to be more important in literacy than in numeracy. This point was also made by the USAID project which pointed out that urban girls in Luganda-speaking areas "*have higher levels of access to English by way of teachers, print and media[45]."*

It may seem to be a truism but older girls get higher scores in the learning tests but this result is important for two reasons. First, it seems to confirm that the learning tests are measuring something that increases with age and could therefore be used to assess progress in literacy and numeracy. Second, it means that any assessment of subgroups should be disaggregated by age in order to detect any impact of belonging to the different subgroups. See "*When I have a problem*" below.

Table 29 - Learning Test Results by Age Set					
	Mean Literacy Mark	Change in Literacy mark	Mean Numeracy Mark	Change in Numeracy Mark	n
6-8	22.8	10.8	16.8	7.4	4
9-11	25.7	9.4	22.2	7.0	80
12-13	28.8	10.7	26.7	7.9	159
14-15	34.9	8.1	33.7	7.4	176
16-17	41.8	8.1	39.4	4.7	142
18-19	41.2	7.1	38.6	3.3	103
20+	42.1	5.2	40.3	2.4	34

The most important finding from the work of analysing the learning test outcomes according to the characteristics of the GEC girls is that so many of the questions in the Household Survey do not lead to the identification of useful subgroups. Useful subgroups would be easy to identify and help the project in providing appropriate support. Subgroups would also help the EET in assessing levels of change attributable to project activities. There seems to

[43] UWEZO (2016): *Are Our Children Learning?* Uwezo Uganda 6th Learning Assessment Report. Kampala: Twaweza East Africa.
[44] USAID/Uganda School Health and Reading Program (2014) The Status of Early Grade Reading and Teaching Reading in Primary School: Cluster 2 Baseline Report, May 2014.
[45] SHRP, Cluster 3, Follow-Up 3, January 2016, page 4.

be no link between performance in the learning tests and most of the variables relating to material wellbeing; the girls' safety; levels of school management; teaching quality; things paid for in girls' education; time spent on household chores; any of the questions relating to Life Skills or decision-making about the girls' lives. It seems that where the question divides the girls in reasonable large subgroups there are no discernible differences in performance in the learning tests. Many of the differences that appear in the data disappear when the data are disaggregated by age. See Table 30 for an example from the Life Skills questions.

Table 30 - When I have a problem – Subgroup Learning Test results BL - ML						
When I have a problem …	Mean Literacy Mark	Change in Literacy mark	Mean Numeracy Mark	Change in Numeracy Mark	n	Ave age
I normally work out how to solve it on my own	39.5	8.2	39.0	6.2	27	17
I ask a friend to help me	38.0	9.4	35.5	5.6	130	15
I ask a trusted adult to help me	34.9	8.3	32.6	6.1	561	14
I do not know how to solve it	16.6	4.1	15.7	5.0	7	12

This table of results looks very clear. The more confident the girls are, the higher their Learning Test marks. The only slightly odd thing is how much progress the least confident girls have made since Baseline. But in fact, this is really a table based on the age of the respondents and this becomes obvious when the data are disaggregated by age (see the right-hand column). It is also important to note that there are only 7 girls in the subgroup of the least confident (and youngest) respondents. This table is presented partly to share the examination of confidence of girls[46] which is a key element of success and the sustainability of success for the project. It is also partly to show how easy it is to find patterns in data where none exist.

3.6.1 Characteristics - Disability

The disability scores in the HHS are based on asking questions on six different forms of difficulty: seeing, hearing, walking, memory, washing/personal care and communication in the girl's first language. Each question can be answered on a scale from 1 to 4: 1= no difficulty; 2= some difficulty; 3= a lot of difficulty and 4= cannot do at all. To identify girls with disabilities we add the scores so a girl with a score of 6 has no disability; a score of 7 means that the girl has some difficulty, and so on. This method is used to identify girls in the database who may have some level of disability. We follow the FM definition of disability as being someone who has recorded a *lot of difficulty* or *cannot do at all* (scoring 3 or 4) in any one domain. This leads to the identification of a subgroup of 23 girls which represents 2% of the sample. Devries *et al* (2014) found 9% of boys and 8% of girls self-identified as having a disability[47]. Barriers to attendance mean it is likely that there are many girls with disabilities who remain out of school which is why the numbers are relatively low.

Table 31 - Frequencies of difficulty scores in survey population		
Difficulty score	Baseline	Midline
6	81%	82%
7	12%	13%
8	5%	3%

[46] See also on confidence Girls on Decision Making under 6.4 Life Skills and The Girls Themselves under section 5. Sustainability Outcome.

[47] WHO https://www.who.int/news-room/fact-sheets/detail/disability-and-health and Action on Disability and Development say 15% - https://www.add.org.uk/

9 or higher	2% (n=24)	2% (n=21)

The EET has informed the project of the findings and alerted the staff to a small number of girls with disabilities (GwD) who were not recontacted for Baseline so that a check can be made on whether their disability is part of the reason.

Disability data from the Household Survey are limited because there are so few girls who are identified as having a disability. The patterns that emerge where a higher level of difficulty coincides with a lower mark and less progress are shown in the table below. We do not assume that any particular difference in marks or in uplift of marks is significant[48]. Comparisons where the differences are smaller than a few points are not shown, for example, difficulties in seeing do not show differences in learning test scores. The learning test scores for girls who do not have a disability are included to show the where a minor difficulty seems to correlate with lower Learning Test scores.

Table 32 - Disability – Subgroup Learning Test results BL - ML

Type and level of difficulty	Mean Literacy Mark	Change in Literacy mark	Mean Numeracy Mark	Change in Numeracy Mark	n
Hearing, no difficulty	35.6	8.4	33.4	6.0	689
Hearing, some difficulty	30.9	8.2	27.3	4.3	29
Walking no difficulty	35.6	8.5	33.2	5.9	709
Walking some difficulty	28.2	5.5	29.8	8.5	13
Walking a lot of difficulty	15.7	8.3	19.5	6.5	2
Remembering, no difficulty	36.0	8.5	33.5	6.0	664
Yes, some difficulty	32.0	8.0	30.5	6.2	52
Yes, a lot of difficulty	14.0	5.1	18.3	2.6	8
Self-care, no difficulty	35.7	8.5	33.3	6.0	718
Yes, some difficulty	12.1	3.3	14.8	2.0	5
Communicating, no difficulty	35.5	8.4	33.2	6.0	719
Yes, some difficulty	33.1	10.9	29.5	5.4	6
Disability[49] prevents you going to school - Never	34.3	8.9	32.1	6.3	107
Yes sometimes	30.3	6.6	30.2	5.5	21
Yes often	16.0	4.2	19.2	5.2	2

The learning test data show some cases where a disability leads to lower scores and smaller progress made between Baseline and Midline. The cases of difficulties in walking and remembering seem strong as the girls with greater difficulties show less good results in the learning tests. The situation might appear clear-cut but the fact that the same does not apply to all the disabilities examined means that some thought is required. Notice for example some very small numbers of girls in the subgroups who have difficulties and the absence of girls with more serious difficulties in some areas.

[48] Nor are we going to carry out T-tests for significant differences between the means that are presented or the other permutations of comparisons that the data would allow.
[49] The respondents to this question are those who have previously identified some level of difficulty.

The girls who have more serious difficulties do perform poorly in the learning tests but it is important to notice that they are making improvements in their learning test scores. These girls may not achieve very highly in literacy or numeracy[50] but they have the possibility to enjoy the other benefits of going to school including increasing levels of confidence and ambition and delaying marriage and their first child.

3.6.2 Characteristics – Wealth

As described in Chapter 1 of this report, the main beneficiaries were selected in such a way as to create a sub-group made up of the poorer families from each community. Within the subgroup we found it possible to identify three subgroups based predominantly on the main source of income of the HoH. This was justified by strong overlaps between this variable and the highest level of education achieved by the HoH. IN fact, it also links loosely to frequency with which the HoH makes savings and the number of meals eaten daily in the household. The groups created are based on: A - having a professional job with a regular salary; B being a farmer with very little education and the rest who form the third group. In this report we have tried to separate the larger undefined group by creating two groups: C - Small business and Skilled artisan from another two groups: D -No earned income and Casual work. Having four groups seems more useful but the differences between the groups C and D do not appear very clearly in some analyses.

Table 33 - Characteristics - Wealth and learning tests marks

Characteristic	Mean Literacy Mark	Change in Literacy mark	Mean Numeracy Mark	Change in Numeracy Mark	n
HoH is Male	35.0	9.1	33.2	6.2	335
HoH is Female	35.9	7.9	33.2	5.7	391
Wellbeing category A	38.9	8.0	36.5	6.2	59
Wellbeing category C	38.0	10.7	35.2	6.5	205
Wellbeing category D	35.4	8.8	31.8	6.0	156
Wellbeing category B	32.9	7.0	31.9	5.4	275
Savings Not at all	34.3	8.8	31.4	5.9	231
Savings Yes, rarely	34.1	7.6	33.0	6.2	110
Savings sometimes.	36.0	7.9	34.1	6.5	199
Savings often	37.0	8.9	34.3	5.1	150
Didn't pay for school materials	33.5	8.5	30.8	6.7	155
Did pay for school materials	35.9	8.9	33.7	6.4	495

There seems to be no difference in Learning Test results of girls where the Head of household is male or female. The wealth category does seem to have an influence although it may be weaker than it appeared in Baseline. There are also some differences between the scores of girls living in households where the HoH saves regularly but this may be a proxy for material wealth. We would expect this characteristic to break down because so many PCGs have joined savings clubs with the support and encouragement of the project. Most households (76%) had to pay for school materials and this seems linked to higher Learning Test results. This too may be an artefact of the level of material wellbeing – the 24% who did not pay for school materials may overlap strongly with wellbeing category B.

These characteristics which directly or indirectly relate to material wealth correspond to the most cited barrier which is financial whether expressed in terms of paying for fees or in terms

[50] Their Learning Test scores are very low and some have learning difficulties.

of lack of money. This has been such a strong refrain in the project reporting and the Baseline Report it seems rather redundant to examine the findings further.

Table 34 - Characteristics – Education and attitudes to education					
Characteristic	Mean Literacy Mark	Change in Literacy mark	Mean Numeracy Mark	Change in Numeracy Mark	n
HoH No school level	33.2	9.6	28.5	5.2	64
Some primary	33.7	7.6	32.4	5.8	270
Primary completed	37.0	8.4	34.4	6.5	125
Completed Secondary 6	38.2	10.6	32.2	2.5	11
After-school vocational training	37.6	10.2	34.8	6.2	25
University completed	43.4	8.5	37.2	7.1	27
How much education should a girl have?					
Some secondary	27.3	5.9	25.9	5.5	15
Complete secondary	31.4	7.0	31.0	5.1	72
Vocational institution	35.0	9.2	33.1	5.4	75
University	36.2	8.7	33.6	6.2	488
The girl should decide	37.2	8.1	35.0	5.8	69

The pattern of the highest level of education achieved by the HoH being loosely linked to Learning Test marks is also shown in the data from the PCGs – the small numbers of girls of those who have had more education do better than the larger numbers of those whose parents and carers have had less education. There does not seem to be much effect on Learning Test marks where the HoH or PCG has had some Secondary schooling but not completed Secondary grade 6.

There are several questions in the Household Survey which attempt to understand the attitudes to girls' education but most of them do not yield useful information. This is sometimes because almost all respondents give the same answer (e.g. everyone says that girls' education is important[51]) or where the answers are better spread but do not seem to link to any other observations. One partial exception is the question of how much education a girl should have. Only one respondent replied that primary education was sufficient and their girl scored very poorly in the learning tests. Despite the very uneven sample sizes there does seem to be pattern with those arguing for more education having girls who score more highly in the learning tests.

The link between highest level of education achieved and attitude to education has not been made explicit except that there seems to be a strong overlap. This may be confounded by the fact that there may also be a strong overlap between material wealth and level of education. Of course, this is a generalisation, and everyone knows someone who is very rich and very poorly educated.

The qualitative work regularly raises questions about parents' attitudes to education. Male parents in the Open Qual [10, 14] for example, talk about "some" parents who have not changed and mothers [11, 12, 15] mention parental "neglect" and some parents who are rarely home or come home very late in the day. The men describe a situation in which parents should not be comfortable if their neighbours have girls who are out of school

[51] Tiny numbers (<20) dissent from the "correct" views on girls' education on questions like: a girl is as likely as a boy to use her education; would you prioritise education for a boy or a girl; it is worth investing in a girl's education; and similar.

because they may be a bad influence on your own children. These reflections focus attention on the role of parents in discouraging their children from attending or doing well at school, perhaps only by being passive or absent.

The question of who decides whether a girl goes to school or not is addressed twice in the Household Survey once with a question to the PCG and then the same question is addressed to the girls. Table 31 shows the results in terms of the percentages who say that the decision is taken by the girl, by the PCG or taken jointly. The results suggest that the parent or carer is deciding alone or with the girl in three-quarters of all cases and the girl alone in one quarter (25%). At Baseline the responses were slightly less consistent and suggested that there must be households where the PCG was saying that the decision making was shared and where the girl was saying that the decision was made by the PCG or another adult. The point is, perhaps rather obviously, that the project must work with the PCG in order to get their support for the girls' education.

Table 35 - Who decides when the girl's education ends?				
		Girl decides	Joint decision	Adult decides
PCG	Control	17%	26%	57%
	Intervention	24%	28%	47%
Girl	Control	26%	20%	54%
	Intervention	32%	22%	46%

The PCG may agree to support the girl's education because they are convinced by argument or by observation and the CLC is an interesting method for allowing a parent to see what happens when a girl starts to make a success of education without incurring any financial cost or taking additional risks.

There is an interesting point in the results in Table 35 which is that the results for treatment and control are significantly different (Girl – p=0.05; PCGs – p=0.02). This may seem surprising as the percentages are not that different but the raw data suggest that Intervention girls are more active in taking decisions about their education than Control girls. This theme is taken up with similar evidence from other decision-making under Life Skills – Girls on Decision Making. where the data from the Household Survey is supported by qualitative interviews to make a case that intervention girls are becoming more confident.

Of course, the age of the girls is important in determining where the power lies in decisions about the girls' lives. Table 36 shows how decision-making moves away from the parent alone to joint decision-making and the girl taking her own decisions. This cross-checking makes us more confident in the data.

Table 36 - Who decides by age set			
Age set	Girl alone	Jointly	Adult alone
9-11	16%	19%	65%
12-13	17%	24%	59%
14-15	19%	28%	52%
16-17	24%	33%	43%
18-19	29%	36%	35%
20+	31%	31%	37%

Subgroups by barriers

We have searched intensively through the results of the Household Survey for the responses to questions that highlight barriers to performance at school that have an impact

on learning test scores. There seem to be very few for the same reasons that the characteristics examined in the survey also fail to deliver clear effects in the learning tests.

Very often the responses are very one-sided and when there are reasonable variations in the responses, there is no relationship with better or worse performance in the tests. Two things do clearly help girls to do well in the tests see Table 37. One is to attend a CLC. The other is to *have an adult (not in your family) who encourages (or encouraged) you to go to school?* Very few girls say that there is no one who encourages them to attend school but they appear to suffer from this situation in terms of their results in the learning tests. For GEC girls this person might be the CRANE Mentor who engages with them and their family on schooling questions.

Table 37 - Barriers – at home and at school

Characteristic	Mean Literacy Mark	Change in Literacy mark	Mean Numeracy Mark	Change in Numeracy Mark	n
Have you ever attended a CLC?					
No	34.3	9.0	32.0	6.6	389
Yes	36.8	7.8	34.5	5.3	337
Is there someone who helps you attend school? No	29.3	7.8	28.1	6.7	75
Yes	36.3	8.6	33.9	5.9	643
I read when I am at home	36.2	8.7	33.8	6.3	652
I never read when I am at home	29.8	6.1	28.9	3.7	67
Does doing chores around the house make you late for school?					
Yes, often makes me late	26.3	7.1	22.3	0.2	20
Yes, sometimes makes me late	27.4	7.7	29.7	4.1	53
No, doesn't make me late	36.4	8.6	33.6	6.2	645
The main way of teaching is for teachers to write on the board and students copy in their exercise books					
Strongly agree	31.9	10.0	29.7	7.0	179
Agree	33.1	8.3	30.8	6.0	184
Disagree	39.1	7.3	37.0	6.8	146
Strongly Disagree	40.5	10.5	37.7	6.6	110

Reading at home, as shown also at Baseline, correlates with better scores in the learning tests. It is not clear what this result means partly because there are so few girls who claim not to read at home. The reasons given for not reading at home do not form a coherent set of responses that would help frame an argument. It may be that not reading at home and doing weakly in learning tests are two symptoms of the same characteristics.

We tested the data from the questions on household chores (Which chores do you carry out? How much time do you spend on household chores?) for effects on learning test scores but there were no obvious patterns. The new question, *Does doing chores make you late for school*[52]? does seem to show a pattern with being late more often links to lower learning test scores.

[52] The earlier questions about how much time girls spend on chores and whether doing chores stops a girl from attending school did not yield any interesting findings.

Similarly, we tested the questions that relate to quality of school management and quality of teaching and found nothing that sheds light on the situation despite the belief that better run schools and better teaching should be linked to better competencies among the girls. The exception is the "chalk and talk" question which asks if copying from the board is the main teaching method. The girls whose answers mean that this is not the case, implying that more imaginative methods are used, have higher scores.

Questions about safety in school and on the journey to and from school do not link to any observable effect on learning test scores. Nevertheless, the qualitative interviews raise significant issues around girls' safety which are important even if the data from the Household Survey do not demonstrate this. Learning Support Teachers [5, 8] report on the new importance felt in school about child protection where until recently safety was "*taken for granted*" but now it seems right to erect a fence around the school. They also talk of parents saying that child protection is better now that it is seen to be more important. The Headteacher [3] remarked that girls now know that they must not agree to take books to the teachers' house. He also talked of positive discipline which was also widely discussed by Learning Support Teachers as bringing down the amount of corporal punishment very significantly even though they and some Headteachers admit that it has not been completely eradicated[53]. No matter how it is described corporal punishment appears as a barrier to attendance and to learning. In the same way that its removal appears to raise confidence and improve rapport between girls and their teachers.

3.7 Learning while at a CLC

In the Baseline Report we presented data from learning tests which were carried out by girls at the beginning and end of their time at a CLC. The data show girls who performed weakly at the beginning of their time in the CLC making huge advances by the time they do the learning tests again at the end of their stay. Results from the Oral Reading Fluency subtask, for example, show girls who scored 20-40 words per minute when they start, were increasing their scores by 30-40 points when they were leaving the CLC. Girls who scored more highly at the start made much less progress, partly, of course, because of the ceiling effect in this task. Those who scored over 60wpm when they started actually record negative progress on average.

The same pattern emerges from data collected from CLCs between Baseline and Midline (see Table 38). The data are possibly less reliable than at Baseline and include, for example, 6 or 7 girls who make negative changes in their reading speeds of around -80 wpm which do not seem likely to genuinely reflect their skills in reading.

Table 38 - Changes in ORF at start and end of girls' time in a CLC

Initial score	Average increase	n
<50	19	17
50-100	39	9
100<	-9	33
All data	6	59

However, the point of this section is to point out the girls who attend a CLC can make huge improvements in their abilities over the six-month period that they receive good teaching and support in a CLC[54]. These changes in performance especially for weaker students are a large part of what makes it possible for them to enter or re-enter mainstream school and make a success of their education. Partly because they have the competencies to do better in class and partly because they have the confidence. Increased confidence may be the more important.

[53] Corroborated by findings in the Household Survey, see under 6.2 School Management.
[54] The results also point out the dangers of overall averages where there are important ceiling effects

4 TRANSITION OUTCOME

Most of the girls in the beneficiary population have the same transition pathways which are based on getting an education. The starting point for many in GEC1 was that they were out of school or likely to drop out of school. For those who had made it into mainstream school their successful transitions were to continue in school and make a success of education. These transitions are expressed in the project slogans of *Get in, Stay in, Transition on; Transition up and Stay Up*. For most beneficiary girls staying in school and moving up a grade constitutes a successful transition. The transition from Primary to Secondary (P7 to S1) is a particularly important achievement.

Other forms of successful transition involve leaving school and starting work or vocational training. This is particularly the case for older girls. There are conditions on what is considered genuine training and genuine working but most girls go into a fairly narrow range of both work and training. There are formal requirements that work pay above the legal minimum wage but it is unlikely that this would be a real constraint.

Moving from out of school into mainstream school can be a successful transition at almost any age. According to the Transitions Pathways (Table 39) a girl moving from school into work or training during secondary school ("Grade 7, 8, 9") would be a successful transfer.

For girls in most of the Creative Learning Centres a successful transition would be into mainstream school and to continue there for at least a year. This may be a harsher test than that of a girl who is already in school continuing for another year. The point is that the CLC girls are mostly those who face more barriers in getting into school and while six months in a CLC is likely to equip a girl to make progress in school it does not necessarily address other barriers like paying fees and other costs or getting the support of adults and so on.

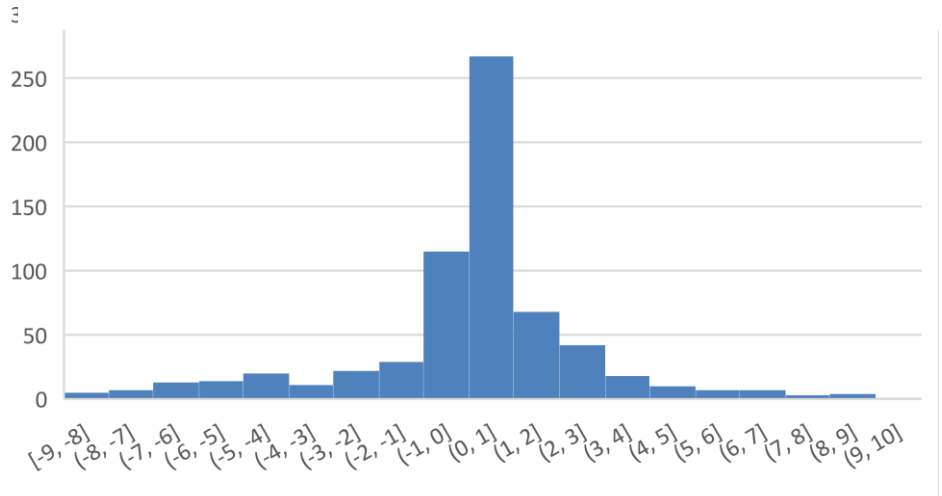
For girls in the two CLCs working with children with disabilities (CwD), their achievements are unlike the transitions of girls in other subgroups and are based on the progress they make in their personal development plans. There is no Control equivalent for these girls. The achievements of these girls are important in their own lives but they cannot be compared with each other or with other transitions. The numbers are small and the differences in rates of transition do not make a meaningful metric. Since there are about 20 girls in each cohort a percentage success rate would not provide any learning. Some girls are too unwell to take part in the activities and will require continuous care; they cannot be assessed as having succeeded or having failed to transition. For these reasons these observations do not form part of the transition calculations for the project. For details on the achievements of the girls in the two specialist CLS, see the section Transition for CLC girls and Table 47 below.

The Transition Pathways table is copied from the Baseline Report and covers the transitions described above. At Baseline the work done on benchmarking did not help very much in fixing targets for transitions. The girls who took part in the survey and their Primary Care Givers (PCG) were asked what the girls had been doing in the previous year and, if in school, what grade they had been in. The situation was made difficult by doing the survey in the first weeks of the first term of a new school year when some of the girls did not know if they were continuing with school and some had not been allocated their grade for the new year.

Table 39 - (Table 10) Transition pathways			
	Baseline point	Successful Transition	Unsuccessful Transition
Lower primary school	Enrolled in Grade 1, 2 ,3	In-school progression Starts new year in the same grade but moves up later in year. Drops out but is enrolled into alternative learning programme Drops out but is recruited to a CLC.	Drops out of school
Upper primary	Enrolled in Grade 4, 5, 6	In-school progression Moves into secondary school	Drops out of school Moves into work, but is below legal age
In CLC		Graduates from CLC and goes into mainstream school. Remains in mainstream school	Leaves CLC and does not enter school or drops out of school within a year of leaving CLC.
Secondary school	Enrolled in Grade 7, 8, 9	In-school progression Enrols into vocational education & training Gainful employment	Drops out of school Moves into employment, but is paid below minimum wage
Out of school	Dropped out	Re-enrol in appropriate grade level in basic education	Remains out of school

At Baseline, the grade to grade transitions were assessed by examining the differences in grades. There was a strong peak of results on and around +1 which signifies a successful transition. There was also a lot of results that could not be interpreted. The data are presented in Figure 7.

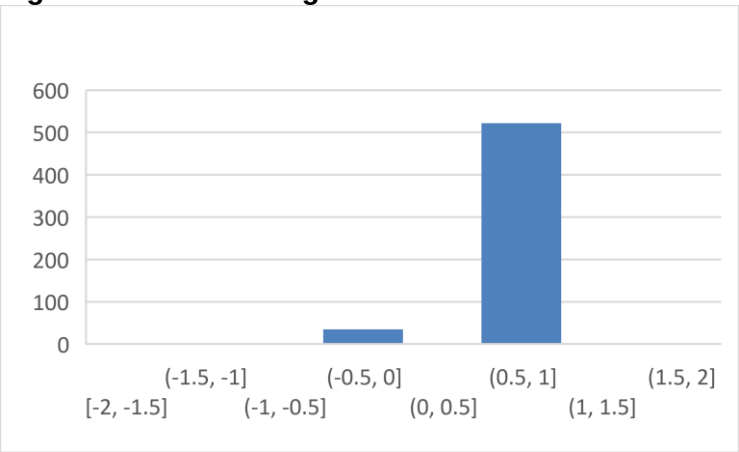
Figure 7 - Grade changes 2017-18



Some of the results may be genuine as girls can move several grades from year to year according to the uncertainties of age and stage and progress made and the availability of spaces in their school. However, some of the reported grade changes are most unlikely to be real and are more likely to be a mistake in reporting or data entry.

The same data for the year change between Baseline and Midline are shown in Figure 8.

Figure 8 - Grade changes 2018-19



The graph may not be the most clear presentation of the changes observed and Table 36 may be an easier way to see how in-school girls have changed grade since Baseline.

Table 40 - Frequencies of Grade Changes at Midline	
Grade change	Frequency
-2	1
-1	1
0	35
+1	522
+2	2

The reason these data are so much neater and accurate than at Baseline is that a huge amount of work was put into cleaning and correction. Grades are notoriously volatile bits of information for a number of reasons[55]. The HHS asks the PCG twice about grades (what grade is girl in now? and what grade was she in last year) and also asks the girl herself the same two questions. In addition, the girl is asked her grade when doing the learning tests. Other helpful information includes the name of the school she attends which sometimes helps to resolve questions of whether the grade is in Primary or Secondary. All these observations from Baseline and Midline were used to assess the reliability of the grades being cited in the surveys. Bringing them all together allowed the correction of errors and the resolution of apparently conflicting grades from different sources and different times. The resulting database provides a clearer and more coherent fix on grades and changes of grades.

The greater accuracy in the data may allow a more strict interpretation of changes at Midline than was used at Baseline where some latitude in either direction was considered acceptable. It seems likely that the identified changes are correct and that one or two girls did not make successful transitions up through the grades.

Benchmarking at Baseline was not considered a successful exercise either for describing the current situation or for setting targets for the future[56]. The sample was probably too small

[55] See also the Baseline Report p1, p31, p35. Some girls at Baseline did not know their grade because term had hardly started and registers had not been finalised. PCGs do not always know the grades of the girls. Girls often do not progress uniformly through the grades but may skip a grade or more and sometimes redo the same grade. The phrase in the HHS "last year" when asked in February may mean the end of the previous academic year or its beginning. Data entry errors are hard to spot with only one year's data – the corrections at Midline became possible because there were more data points that exposed errors or confirmed uncertain entries.
[56] See Baseline Report p65.

and not representative of the GEC girls. Almost all the girls in the Benchmark sample moved through successful transitions by staying in school and the high proportions (some sections recorded 100%) did not leave any room for improvement in future work.

In-School Girls transition rates and targets

The overall rates for transition of in-school girls are very high and in most cases over 90% have transitioned into the next grade. If we adopt the more generous interpretation used at Baseline, that to remain in the same grade (effectively repeating a year) should also be considered a success given the inaccuracies in attribution of grades, the success rates would be even higher. This creates a problem in setting targets since the differences in few percentage points depends on the behaviour of one or two girls.

The pattern of transition for ISG in Control is remarkably stable – all girls move into the next grade except for 5 (3%) who appear to repeat a year. This equates to a successful transition rate of 97% or 100%.

There are interesting differences between the patterns of transition of Control and Intervention girls. Table 41 shows a fairly orderly progression through school grades being made between Baseline and Midline – for example: the 22 girls who were in P3 at Baseline are in P4 at Midline. It seems harsh to reduce the transition rate for the small numbers who remain in school but repeat a year, like the 3 girls who were in P4 at Baseline and at Midline. Overall, it seems fairer to say that the in-school girls in Control are transitioning at 100%. Of the 18 girls who were Out of School at Baseline, 6 have started training and 12 remain OoS making a nominal transition rate of 33%.

Table 41 – Control girls changes between Baseline and Midline[57]												
	Midline Grade											
Baseline Grade	**OoS**	**P4**	**P5**	**P6**	**P7**	**S1**	**S2**	**S3**	**S4**	**TVET**	**Total**	**Transition**
OoS	12									6	18	33%
P3		22									22	100%
P4		3	29								32	91%
P5			1	29							30	97%
P6				2	22						24	92%
P7						5					5	100%
S1							27				27	100%
S2								18			18	100%
S3									15		15	100%
Totals	**12**	**25**	**30**	**31**	**22**	**5**	**27**	**18**	**15**	**6**	**191**	

The transitions of the Intervention In-School Girls are similar to those of Control girls and the majority make their way to the next grade from Baseline to Midline. But an important number (19) drop out after S4 and another 7 girls drop out from grades P5 to S1. Stopping schooling after S4 is a common phenomenon and it is a pity we do not have data on P5 girls in the Control sample. Currently we consider dropping out after S4 to be a failed transition although girls are above the legal leaving age in most countries by that time. Another 19 girls leave school at this transition and start training while one girl goes into a CLC.

[57] Data in this and the following table from variables “M1 Grade FINAL (abbrv)” and “B Grade MOST FINAL”

That seven girls drop out over four grade transitions may not be important for the overall statistics but it could be another minor observation to add to the suggestion that the Control girls are better-off than Intervention girls[58].

Another difference between intervention and Control is the transition from P6 into P7 where 17 Intervention girls do not drop out but resit P6. Note though that transition from P7 to S1 which is normally regarded as a difficult transition, is relatively successful. Three out of 61 girls drop out at that stage but the overall success rate is 95% as others who do not go on to Secondary go into a CLC or into training.

The transitions of girls who were OoS at Baseline are at face value better in Intervention than in Control although the disparity in overall numbers may make the comparison meaningless. Four girls have gone on to university so their OoS status at Baseline may have been due to delays in admission processes. 13 girls have gone back to school and some of these may only have been out of school at Baseline while waiting to be allocated a school place. Nearly half (50/106) of the OoS have started training of some sort and this must be due in part to the activities of the project and the good links that CRANE has with training centres. 34girls remain OoS and are recorded as having a failed transition.

Table 42 - Intervention Girls changes between Baseline and Midline																
	Midline															
Baseline Grade	**CLC**	**OoS**	**P4**	**P5**	**P6**	**P7**	**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**TVET**	**Uni**	**Total**	**Transition rate**
OoS	5	34	3	1		2		3	1	1		2	50	4	106	68%
P3		1	26										1		28	96%
P4			3	45	1										49	94%
P5	1	1		2	52										56	95%
P6	1	2			17	48									68	72%
P7	2	3					53	1					2		61	95%
S1		1					1	28					2		32	94%
S2	1							1	31						33	97%
S3										34			2		36	100%
S4	1	19								2	10		19		51	41%
S5												5			5	100%
S6		4											1	1	6	33%
	11	**65**	**32**	**48**	**70**	**50**	**54**	**33**	**32**	**37**	**10**	**7**	**75**	**5**	**531**	

Transition of Out of School Girls

The PCGs and the girls give slightly different answers to the questions on Out of School girls (OoS). Both are asked if the girl is in school and what the OoS were doing last year. Table 41 gives the responses from the PCGs and shows that most girls who are OoS at Midline were also OoS at Baseline. Most of these girls appear to have made successful transitions, that is, they are in work or training. It is not clear what "Other" means in this context and the question in the HHS needs to be changed for future surveys.

At face value this is an area where Intervention girls are doing better than Control girls with 68% completing successful transitions compared with 33%.

[58] See Annex 3 for discussion of the goodness of match between Intervention and Control samples.

Table 43 - What were OoS girls doing last year?														
	OoS	P1	P2	P3	P4	P5	P6	P7	S1	S2	S3	S4	S5	S6
a) Vocational Training	18											3		
b) In employment	15											2		
c) In domestic activity	14	1					1					1		
d) Other	3			1		1	1	3	1			12		4

Table 42 also shows how those who have left school since Baseline were mostly in S4 and S6. The EET has been in touch with the project concerning the younger intervention girls who have left school and provided the reasons given. The project staff have followed up with the families and two of the three younger girls are back in school at the time of writing (3 months after the girls were identified as having dropped out). The other girl may not go back into school and wants to join the vocational school to learn a skill.

Three girls appear to have failed to transition from P7 into secondary school. The three are in the Intervention group of 58 girls. There are only 5 girls in Control in P7 and they all appear to have gone on into secondary school.

The data on OoS need to be examined with care. The PCGs say in answer to one question that 85 girls are not going to school. Further questions follow on what the girls are doing and under the option "Other" fifteen of these girls are found to be in school or in education. One parent correctly answered that her girl is not going to school because she has just started at university.

Treatment vs Control

There are relatively few Out of School girls in the project and the Control group is smaller than the intervention group so the total number of Out of School girls in Control is small. When the numbers are looked at in terms of the girls who left school since Baseline there appear to be no girls below 18 in the Control group

Table 44 - No young Control girls dropped out of school since Baseline										
Age set	12-13		14-15		16-17		18-19		20+	
T or C	T	C	T	C	T	C	T	C	T	C
a) Vocational Training			1				10	1	3	
b) In employment							3	1	4	
c) In domestic activity	2		2		4		5	4		
d) Other	1		5		6		9		4	1

Although the number of Intervention girls below 18 who dropped out is small it is significantly larger than the zero among the Control girls. This may be a symptom of the weaker educational status of the Intervention girls. this possibility is a major concern for project staff who know that the GEC girls were selected because their educational history was weak.

Transition by Marriage

In the Control population, one Out of School girl got married between Baseline and Midline. In the Treatment group, four in-school girls and one OoS got married over the same period. This seems to imply that getting married had no impact on the transition status of the girls involved – the out of school remained out of school and the in-school remained in school. The age set data suggest that of the four GEC girls who got married, 2 were 14-15 years old and 2 were 16-17 years old. The OoS girl was over 20. The Control girl who got married was 18-19. CRANE staff have suggested that the younger girls are not married but have

interpreted the question to be about being engaged or having an idea of who they will marry perhaps someone with whom they are already in a relationship.

Transition for Girls with Babies

At Midline, about 2% of Control girls and 9% of Intervention girls had a baby that they were caring for according to those who answered the specific question, *Do you have children/babies of your own now*? (n=890). The difference is highly significant (p=0.005). It seems possible that this difference is a feature of the recruitment to the GEC by the Mentors based on their vulnerability. This may not be the same as the recruitment to the control sample by the teachers and head teachers assisted by the EET in 2018. It is probable that the girls offered to the project as control girls were not among the more vulnerable in their communities. There may even have been bias in the other direction, towards stronger students, in the identification of the girls in school.

Table 45 - Age range of babies of Control and Treatment girls

	Less than one year		One to two years		Over two years	
Age set	Control	Treatment	Control	Treatment	Control	Treatment
14-15		1				
16-17		1		3		
18-19	1	5	1	5		9
20+		1		3		4
Totals	1	8	1	11	0	13

Unfortunately, we don't have entirely reliable data on all those who said they had a baby in order to fix their age set and their in-school or Out of School status. This is why the data in Table 44 do not exactly match the data in Table 45.

Table 46 - Age range of babies of ISG and OoS girls

	Less than one year		One to two years		Over two years		Totals	
Age set	Control	Treatment	Control	Treatment	Control	Treatment	Control	Treat
ISG		6	1	8		14	1	28
OoS	2	5	1	4	1	10	4	19
Totals	2	11	2	12	1	24	5	47

In terms of transitions there do not appear to be significant numbers of project girls in the sample who are making unsuccessful transitions by having a baby. It looks as if eight GEC girls had a baby in the time between Baseline and Midline surveys. Of those, six appear to be in school (there are 3 more babies to account for in the data relating to being in-school and it only seems possible that these were born to girls over 20 who did not provide accurate age data). Two girls in the Control sample had a baby over the same time period and both of these were out of school at the time of the Midline survey.

Early marriage and having babies are frequently mentioned as a reason for girls dropping out of school and this is undoubtedly the case in a general sense. The Household Survey data tell a slightly different story. First, the issues of early marriage and early childbirth are relatively infrequent in the sample but much more prevalent in Intervention than in Control. If this is a real phenomenon, it is probably due to higher levels of vulnerability in the Intervention sample because of the ways the GEC girls were recruited. Second, getting married and having a baby seem to have less effect on girls dropping out of school in Intervention than in Control. The numbers are so small that these comments include some conjecture which requires cross-checking through interviews with the girls concerned or with CRANE staff who know the girls' situations. It is possible that the numbers do not have an

impact on transition calculations because the project is working at keeping the GEC school-age mothers in education.

Transition for CLC girls
There should be a new table within the Outcomes Spreadsheet for the CLC girls as their stories of transition do not fit with the current settings.

Table 47 – CLC girls activities at Midline[59]																
	No data	**OoS**	**CLC**	**P1**	**P2**	**P3**	**P4**	**P5**	**P6**	**P7**	**S1**	**S2**	**S3**	**S4**	**TVET**	**Total**
CLC	31	14	4	1	2	8	4	5	3	4	1	5	2	1	20	105

Notice also from the previous table 42 that 6 girls transitioned in to a CLC from grades P5 to S4. Older girls leaving school to join a CLC is a relatively new phenomenon. We have had interesting interviews with two of these girls – one who left after S1 and the other from P5. In both cases, this seems to have been good for the girls involved. One went from the CLC on to training in hairdressing and the other was still in the CLC and training in tailoring/dressmaking [OQ2 and OQ3].

Also in Table 42, the more normal recruitment to CLCs is shown by the transitions of 5 OoS girls in to a CLC.

There are 404 girls in the Midline sample who said that they have attended a CLC. The data from the girls and their PCGs are slightly different but tell roughly the same story that almost all of them are in school and a small number may not be in successful transitions.

Table 48 - What CLC girls who are out of school are doing					
Source	Vocational training	Employment	Domestic work	Other	Total
Girls	16	11	7	18	52[60]
PCG	18	9	12	15	54

It is unfortunate that the Household Survey allowed the vagueness of the answers Domestic work and Other. Domestic work is a common way of making money and it may be a chosen economic activity. The task of washing clothes could earn a girl UGX15,000 a day; a full-time housemaid might make UGX150,000 per month. Domestic work could be an imposed activity to help out with income or managing the home.

If we assume the worst case that all girls in Domestic work and Other are working in unproductive and poorly paid or unpaid positions, we would conclude that these were unsuccessful transitions at a rate of 7% of all CLC girls. If we assume that only the girls doing "Other" have not been successful, the failure rate drops to 3%. That would mean that 97% of CLC girls are currently in a position which represents a successful transition.

Studying the CLC girls' progress and recording the observations need to be improved. The girls in the CLCs are sometimes confounded with OoS girls (because technically they are not in a school) and the questions relating to what the girl was doing last year do not allow the option that the girl was in a CLC for part of the year and in school or OoS for part of the year although those options are possible. Data for CLC girls should be collected as a separate category which will involve some modifications to questions and to the routing of the

[59] It was not possible to do the analysis of CLC girls' transitions using the same data sources as the other two tables (Control girls and Intervention girls). CLC attendance is recorded in variable "B Grade Abbr (abbr)"
[60] The girls did not include 2 Midline girls who say they are currently in a CLC.

Household Survey. The CLC girls should form a specific study at least on some aspects of learning and transition.

Reasons why girls drop out

The Household Survey asks three separate questions on the reasons girls are not in school. First, it asks the PCG if it is likely that the girl will be in school next term and follows up on those who say that it is unlikely with a request for the reasons. The responses are shown in Table 49. The respondents can give as many answers as they choose. The main finding is, as always, the question of cost. This applies to all the questions about barriers to attendance. The second most common reason is that the girl has finished schooling and this is also shown by the age of the OoS.

Table 49 - Why is girl unlikely to be in school next term

Reason	Count
Costs too high	60
Already graduated	14
Girl has to look out for family members	8
Girl got married	7
Girl is not interested	4
The girl has had enough schooling	3
The girl is too old to attend school	1
Unsafe to travel to school	1
Girl was weak in studies	0
School is not safe	0
School does not have special facilities to help the girl	0
Poor infrastructure at school (girls have problems using toilets at school, access to classroom, getting around school,)	0

A second question on reasons for being out of school OoS is addressed to the PCGs who have said that their girl is not in school. The results are in Table 50 and again financial concerns cover the majority of the responses. Concerns over the journey to school cover the next most commonly cited reasons. There are more concerns over safety on the journey than there are about safety at school which corresponds to the findings of the Ministry of Gender and Devries et al[61]. The problem of paying for schooling is the only clear common ground between the two different attempts to find out why girls may not be in school. The other reasons mentioned seem to attract different rates of response.

Table 50 - Why is girl not in school this term

Reason	Control count	Intervention count	Totals
a. Was there enough money to pay for the girl's schooling?	8	59	67
b. Does the girl need to work, earn money?	6	48	54
e. Is the school that the girl would attend too far ?	3	34	37
c. Is it unsafe for the girl to travel to/from school ?	3	21	24
h. Death of significant family member(s)?	2	15	17
k. The girl failed exams?	1	15	16
i. Illness of family member(s)?	2	11	13
g. Is the girl in vacation (Awaiting results)?	0	10	10
d. Is it unsafe for the girl to be in school ?	0	4	4
f. Does the girl need special services which are not available?	0	3	3
j. Domestic violence in the home?	0	2	2

[61] Both references are cited earlier.

The third attempt is addressed to the Out of School girls themselves and again financial costs are by the far the most frequently cited. The questions are presented in both negative and positive versions in the Household Survey[62] so the responses in Table 51 may appear illogical. For example – Was it safe for you to be in school? 72% said they felt safe so the 28% in the table is those who did not feel safe and for whom this might have been a contributory barrier to them staying in education.

Table 51 - Why girls are out of school – the girls' reasons (% of girls who mention)	
Reason	Percent
OS1. Was there enough money to pay for your schooling?	77%
OS2. Did you need to work, earn money?	45%
OS5. Was the school too far?	39%
OS6. Did you have to help out at home?	37%
OS16. Do you have a baby or a child?	29%
OS4. Was it safe for you to be in school?	28%
OS12. Was there a toilet you could use easily at school?	27%
OS3. Was it unsafe to travel to/from school?	25%
OS7. Did you need special services or assistance to attend school?	24%
OS20. Were you bullied by other pupils?	17%
OS15. Are you married or about to get married?	13%
OS10. Were you refused entry into the school?	12%
OS23. Did you stop going to school because of Illness of a family member(s)?	11%
OS9. Did teachers mistreat you at school?	9%
OS22. Did you stop going to school because of the death of a significant family member(s) ?	9%
OS25. Did you stop going to school because you failed exams?	9%
OS14. Have you done enough schooling?	6%
OS8. Did you stop going to school because assistive devices such as braille textbook, hearing aid, glasses, wheelchair were not available?	5%
OS13. Are you too old/young to attend school?	5%
OS11. Was it difficult for you to move around the school?	4%
OS24. Did you stop going to school because of the experience of domestic violence	3%

It is important to note three other possible responses to this question which show that the girls did not appear to drop out because of a lack of interest in education.

Table 52 - Out of school girls say school is important	
Was going to school important for you?	89%
In your opinion, does school help girls get a job?	88%
Are/were you interested in going to school?	76%

Transition for Girls with Disabilities

There are two CLCs which specialise in working with children with disabilities. All the children are assessed individually and assigned a personal development plan. Some have very severe impairments and require continuous care and support. Others are able to learn skills in crafts and start making scarves, mats, bags and jewellery to sell. Others learn to run small enterprises in production or small-scale marketing. A small number of girls graduate

[62] A recognised technique to check for acquiescence effects.

from these CLCs and join a mainstream school or a special school. The current activities of the latest cohort of 22 girls in the two CLCs are shown in Table 53.

Table 53 - What those with severe impairments did next	
Current activity	Number
Continuing in the CLC	3
Moved to vocational school learning craft skills	4
In mainstream school	3
Graduated from vocational school and set up enterprise	7
In special school	2
Too unwell to work – mother running small enterprise	2
Lost contact	1

At the time of writing 16 of the 21 with whom the project has been in touch have made successful transitions. If the three who are still in a CLC are added, since they haven't actually failed to transition at this stage, the success rate would be increased. The two girls who are too unwell to work cannot be assessed as having succeeded or having failed to transition.

It is a cause for some celebration when a girl leaves the CLC for a mainstream school but it's not clear that this is a more successful transition than entering a specialist school for people with a disability or managing pig or goat-rearing, running a bakery of selling craft items. The CLC staff work closely with the parents and develop plans with them and also practice savings and loans management. A great deal of the success of the transition depends upon the work of the parents. Talking to parents of children at these CLCs makes one wonder if a major part of the transition is in the minds of the parents. They tell stories of how they and their neighbours changed their ideas about what the child can achieve and how their appreciation of the child is transformed[63].

CLC staff also mention the value of breakfast meetings with parents and children where it is possible to interact on an informal basis and exchange ideas easily. This provides some time to talk about issues which is not under pressure while the children are eating or being fed and helps parents see other ways to understand disability.

Integration of girls with disabilities into mainstream school is also helping according to qual interviews with CRANE staff. There are a number of significant successes where girls with disabilities are studying alongside girls without disabilities and this helps to reduce the stigma associated with disability[64]. One disabled girl was offered a trial in a mainstream school and then turned down the invitation to return to the CLC where she had been attending.

[63] Qual interviews 1-20 at Baseline – Baseline Report.

[64] New Hope School is cited as an example where 4 girls with disabilities are learning together with other girls. This observation could be put in the section on Attendance, Inclusion, Teaching Quality, School Management or here in Transition.

The following table is mandated by the Fund Manager and forms part of the Outcomes Spreadsheet. The samples sizes are included. No useful observations can be extracted from these data. Ages are the ages of the girls at Baseline.

Table 54 - Transition success rates

	Intervention		Control	
Age	Transition success rate	n	Transition success rate	n
6		0	100%	1
7	100%	6	100%	5
8	100%	14	78%	9
9	94%	16	95%	22
10	97%	37	100%	25
11	93%	61	87%	23
12	90%	82	100%	23
13	90%	68	100%	30
14	93%	67	88%	16
15	89%	57	95%	19
16	89%	57	80%	10
17	84%	50	44%	9
18	78%	41	67%	6
19	74%	42	50%	2
20	58%	12	33%	3
21	71%	7	0%	1
22	40%	5	100%	1
23	50%	4	0%	2
24		0	0%	1
25	0%	1		0
26				
27				
28				
29	0%	1		0
30		0	100%	1

Target setting for transition

The Outcome Spreadsheet by its own hidden mechanisms suggested improvements in Transition of 5, 7 and 8% over the post-Baseline evaluation events. The project suggested 2, 3 and 4% instead given the already high levels of transition recorded in the Benchmarking and the Baseline report. The percentages in the benchmarking by age vary from 100% to 33% and the eight age-set values that seem to relate to actual transitions have an average of 93%. It is hard to see how a project could be expected to improve on such a figure on an annual basis.

Tables 43 and 44 examine the numbers of girls of school-age who have stopped attending school since Baseline. If we take the worst case and assume that all those who are in Domestic work and Other are in fact doing nothing, the transition rate for in-school girls would be 97%[65]. The practical point from these observations and calculations is that the

[65] Assuming all 20 girls below 18 years old are drop-outs and the school population within the sample is 603.

project is doing well in supporting girls to stay in school and in following up on those who stop attending. The numbers are small and become smaller when some of the girls who dropped out return to school and others start gainful activities.

It may be best to develop new targets according to the different transitions disaggregated by the grade changes or the type of transition. Most grade to grade transitions of In School Girls cannot be improved upon. This is partly because the transition rates are already very high and partly because the numbers in each grade are relatively small. The difference of one or two girls succeeding or failing at the time of the survey would make more than 4 or 5% difference in the rates.

It would be reasonable to aspire to the following targets:
Grade to grade transitions which are already over 90% should stay above 90%.
Over 90% of CLC girls should complete a successful transition in the year after they leave a CLC.
The overall success rate for CLC graduates should be above 80%.
Successful Transitions from OoS and from P6 to P7 should improve by 5%.
Transition from S6 should improve by 5%
Transition from S4 should improve by 5%

5 SUSTAINABILITY OUTCOME

The sustainability indicators of the project are divided into four areas rather than the three specified in "Table 13" see Table 55. The community work is separated into two areas covering the work at family level and the work at community level. The project is putting different types and amounts of work into these two areas and expects to find results changing in different ways.

Table 55 shows overall good progress according to the scoring of the indicators by the Mentors who work in seventeen different locations. The scoring exercise is part of their monitoring work and is facilitated by members of the EET. The exercise was carried out in January 2019 and September 2019.

Table 55 – (Table 13) Sustainability indicators

	Community Family	Community Leaders	School	System
1	Material support to girls' ed	Engagement in schools	Engagement with parents	Local
2	Involvement in schools	Moral support to girls' education	Engagement with students	District
3	Moral support for education	Child Protection	Teaching methods	National
4	Child Protection		Child protection	
5			School administration	
6			Management of teachers	
7			Inclusion of CwD	
Baseline Sustainability Score (0-4)	**2.22**	**2.43**	**2.56**	**1.76**

Overall Sustainability Score (0-4, average of the three level scores)	**2.24[66] (2.30)**			
Midline sustainability Target (0-4)	**2.42**	**2.63**	**2.76**	**2.16**
Midline score (0-4)	**2.47**	**2.65**	**2.99**	**2.53**
Overall sustainability Score (0-4, average of the three level scores)	**2.66 (2.73)**			

The learning and the usefulness of the monitoring of the Sustainability Scorecard comes from the discussions around the scores that are given by the Mentors during and after the exercise.

Table 56 - Sustainability Scorecard Results Community - Family

	Community - Family											
	IND 1: Material support to girls' education			IND 2: Involvement in schools			IND 3: Moral support for education			IND 4: Child protection		
Location	2018	2019		2018	2019		2018	2019		2018	2019	
1	4	2	-2	2	2	0	2	1	-1	3	2	-1
2	3	2	-1	3	1	-2	2	3	1	2	2	0
3	3	3	0	3	3	0	2	4	2	2	4	2
4	3	2	-1	2	2	0	2	2	0	3	2	-1
5	2	3	1	2	2	0	1	3	2	3	4	1
6	2	3	1	3	2	-1	3	3	0	2	3	1
7	2	3	1	1	3	2	2	3	1	2	3	1
8	2	3	1	1	1	0	2	3	1	3	3	0
9	2	3	1	2	4	2	2	3	1	2	3	1
10	2	3	1	1	2	1	2	4	2	3	1	-2
11	2	2	0	3	3	0	2	3	1	2	3	1
12	2	1	-1	1	4	3	2	2	0	2	2	0
13	2	2	0	3	2	-1	4	3	-1	4	1	-3
14	2	3	1	1	2	1	3	2	-1	3	3	0
15	2	2	0	1	2	1	2	2	0	2	1	-1
16	2	2	0	1	2	1	2	2	0	3	3	0
17	2	2	0	2	3	1	2	2	0	2	2	0
Totals	**39**	**41**	**2**	**32**	**40**	**8**	**37**	**45**	**8**	**43**	**42**	**-1**

The Mentors' scores on Child Protection require special attention. Taken at face value they appear to suggest that it is an area of work with families and leaders that has gone rather

[66] The first average is calculated by giving equal weight to the four areas of project activity. The average in brackets is the average calculated by giving equal weight to each indicator. It makes more sense to assess progress by giving equal weight to each area of project activity but the greatest learning comes from discussions about the scoring of each indicator and interpreting the composite scores.

poorly. Work with schools appears to have made more progress but rather less than in other indicators of change at School level. The Mentors explain their lower scores at Midline in terms of their lower understanding of child protection issues at Baseline. Raised awareness leads to lower scores – this is a case of the "crime statistics" problem (see footnote 47). In fact, a great deal of work has been done on child protection with parents and teachers. Mentors have an additional problem with this area of the Sustainability Scorecard as it was not clear to them if they are scoring the progress made in child protection with the GEC families or with the community more widely. They would, not surprisingly, give better scores for work with the GEC families and see less progress at the level of the community (see Table 56).

The Sustainability Scorecard provides interesting insights into the geographical variations of progress. Partners in location 1 do not manage schools of their own in the way that other partners do but provide support to schools which makes their work less direct. They report more difficult relationships recently with schools and give lower scores in the Sustainability Scorecard exercise. In contrast, partners in location 10, despite negative scores in child protection, mostly for the reasons given in the preceding paragraph, are very positive about progress being made recently. This is the partner working Karamoja where they are able to see important progress made in a short time[67]. The partners report establishing school clubs for the first time and benefiting from very good relationships with district and local leaders.

Table 57 - Sustainability Scorecard Results Community - Leaders

	IND 5: Engagement with schools			IND 6: Moral support to girls' education			IND 7: Child protection		
	2018	2019		2018	2019		2018	2019	
1	1	3	2	2	2	0	3	3	0
2	2	3	1	3	2	-1	3	2	-1
3	2	3	1	1	3	2	1	4	3
4	3	2	-1	2	2	0	3	2	-1
5	4	4	0	4	4	0	3	3	0
6	1	3	2	2	2	0	3	1	-2
7	4	3	-1	4	3	-1	3	3	0
8	3	2	-1	3	2	-1	4	3	-1
9	1	2	1	1	2	1	2	2	0
10	2	3	1	1	3	2	3	2	-1
11	2	2	0	2	2	0	1	3	2
12	2	3	1	3	4	1	3	4	1
13	3	2	-1	2	2	0	2	3	1
14	2	2	0	2	4	2	3	4	1
15	1	2	1	3	2	-1	3	2	-1
16	3	4	1	2	3	1	3	2	-1
17	2	2	0	2	3	1	2	2	0
	38	**45**	**7**	**39**	**45**	**6**	**45**	**45**	**0**

Work with community leaders has been seen by Mentors to be a more difficult area of work partly because the project does not see Leaders as direct beneficiaries and does not provide support to them[68]. Also, many Leaders are older men who tend to be less supportive of girls'

[67] See above "New Activities" p4
[68] Of course, some leaders are also parents or members of the savings group.

education and other project initiatives. Community leaders are elected and are replaced regularly and this situation has led to Mentors facing two different problems in reporting on progress in changing of attitudes. First, there are Mentors in areas where the local leaders have recently taken up office and have not attended any project training or events. Second, other Mentors report current Leaders not wanting to be seen advancing unpopular ideas for fear of losing votes. This situation neatly reveals the differences between changing attitudes of people with whom the project has close contact and changing social norms which are maintained in place by much larger numbers of people who have not had any contact with the project.

Work in schools is being made difficult for Mentors because the GEC girls are moving to more different schools and there are now significant numbers in schools where the project is not working. While the Mentors see important changes happening in project schools, they recognise that these changes do not benefit all the GEC girls.

Table 58 - Sustainability Scorecard - Schools																					
	IND 8: Engagement with parents			IND 9: Engagement with students			IND 10: Teaching methods			IND 11: Child protection			IND 12: School administration			IND 13 Management of teachers			IND 14: Inclusion of CWD		
Location	2018	2019	Difference	2018	2019	Difference	2018	2019	Difference	2018	2019	Difference	2018	2019	Difference	2018	2019	Difference	2018	2019	Difference
1	3	2	-1	4	3	-1	2			4	3	-1	4	1	-3	4	2	-2	0	0	0
2	2	4	2	3	3	0	3	4	1	2	4	2	2	3	1	2	3	1	3	2	-1
3	3	3	0	3	3	0	3	2	-1	2	3	1	2	2	0	3	3	0	1	2	1
4	3	4	1	3	4	1	4	2	-2	4	4	0	3	2	-1	4	2	-2	3	3	0
5	3	4	1	3	4	1	3	3	0	2	4	2	2	3	1	3	3	0	2	3	1
6	3	4	1	4	4	0	4	4	0	3	4	1	2	3	1	3	3	0	1	3	2
7	4	4	0	4	4	0	2	4	2	3	3	0	2	4	2	3	4	1	3	3	0
8	2	4	2	4	4	0	1	2	1	3	3	0	2	2	0	2	2	0	1	2	1
9	3	4	1	3	4	1	3	3	0	3	3	0	2	2	0	3	3	0	4	2	-2
10	4	4	0	3	4	1	2	4	2	3	2	-1	3	4	1	3	4	1	1	3	2
11	3	4	1	4	4	0	3	4	1	3	3	0	3	4	1	3	4	1	1	3	2
12	3	3	0	4	4	0	4	4	0	2	2	0	2	3	1	2	4	2	1	3	2
13	3	3	0	4	3	-1	3	4	1	3	2	-1	1	2	1	1	3	2	1	3	2
14	2	2	0	1	4	3	2	4	2	1	3	2	1	4	3	2	4	2	0	2	2
15	3	3	0	3	2	-1	2	2	0	3	2	-1	2	1	-1	2	2	0	2	1	-1
16	3	3	0	3	4	1	3	4	1	2	3	1	3	4	1	3	4	1	2	1	-1
17	2	2	0	2	2	0	2	2	0	2	2	0	2	3	1	2	2	0	2	2	0
Total	49	57	8	55	60	5	44	52	8	45	50	5	38	47	9	45	52	7	28	38	10

Engagement between schools and parents is seen to be improving from the results both in schools and with Parents. These changes are only weakly observed in the Household Survey[69]. This is partly because results were very positive at Baseline and there is little room for improvement. For example – at Midline 94% of PCG in Control schools and 97% in Intervention areas agree that teachers cooperate well with parents. However, a more open question on whether access to school has improved in the last year does show a borderline significant difference between Intervention (77% agree) and Control areas (68% agree) (p=0.10).

Improvements in teaching methods noted by the Mentors are corroborated by the Household Survey and the qualitative interviews as noted in Chapter 6.3. In this chapter we report on differences between Intervention and Control schools where there are significant ($p<0.05$) and almost significant differences in teachers' assumptions about the competencies of boys and girls; the teaching methods used[70], the frequency of receiving information; the frequencies of visits by PCG to school and the overall change in quality of teaching in the last year. Chapter 6.3 also gives reports from qualitative interviews with LSTs and Headteachers who mention different project interventions which corroborate the household survey results and the Sustainability Scorecard exercise.

	Table 59 - Sustainability Scorecard – System								
	IND 15: Local			IND 16: District			IND 17: National		
Locations	2018	2019		2018	2019		2018	2019	
1	1	1	0	2	1	-1	3	2	-1
2	2	2	0	2	2	0	1	3	2
3	2	3	1	2	2	0	1	4	3
4	3	2	-1	2	2	0	2	2	0
5	3	3	0	2	3	1	3	3	0
6	2	2	0	1	1	0	1	1	0
7	2	3	1	1	3	2	1	4	3
8	2	2	0	2	2	0	1	1	0
9	1	2	1	2	3	1	1	3	2
10	2	4	2	2	4	2	1	4	3
11	2	3	1	1	3	2	2	4	2
12	2	3	1	2	4	2	2	3	1
13	2	1	-1	2	1	-1	2	2	0
14	2	4	2	1	2	1	2	4	2
15	1	1	0	1	2	1	2	2	0
16	1	3	2	1	3	2	2	3	1
17	2	2	0	2	3	1	2	2	0
Totals	**32**	**41**	**9**	**28**	**41**	**13**	**29**	**47**	**18**

Staff in location 1 remain pessimistic about the changes that have taken place but all the other Mentors have made significant upgrades in their scores for National level showing their belief in the progress made at this level.

[69] Relevant questions cover how often the PCG visits the school; how frequently the parent is informed of their child's progress; PCGs' awareness of changes in teaching practice and a more general question on how well teachers cooperate with parents.
[70] The "chalk and talk" question.

The interviews with CRANE staff more directly involved with the national level work are similarly very positive. The fact that the Ministry of Education has adopted and endorsed the methodology proposed by the project for the identification of children with special educational needs (SEN) is a huge result for the project. Not all the credit goes to the GEC project since the work builds on CRANE's previous work and reputation and its work as convenor of a working group on SEN. The SEN unit within the Ministry (Department of Special Needs and Inclusive Education (DSNIE) has been strengthened with more staff and a larger operating budget.

CRANE retains a key role in the task force that draws together staff from Department of Special Needs and Inclusive Education within the Ministry of Education and Sports and a range of national and international NGOs who focus on issues of disability and education needs. Staff from these institutions met for the first National Inclusive Education Dialogue, an event which formally launched the National Learning Needs Identification Tool (NLNIT). Also in this event the Ministry committed to a range of activities including the finalisation of a National Inclusive Education Policy and the rolling out of the NLNIT including training for teachers in post and those in training. More training materials will be produced and a national study will be carried out to assess the costs of providing support to children with special needs. The education management systems will be revised to create more complete monitoring of the on different categories of learners.

CRANE will continue to play its convening role in the task force and will support the Kampala Capital City Authority Special Needs Assessment Centre through renovation of the premises and some staff costs. CRANE will also develop the training of trainers course for MoES staff who will provide training to teachers. It is worth noting that the activities are well targeted on the difficulties described by Kristensen *et al*[71] including the weaknesses in identification of special needs and the quality and quantity of materials available in special schools and mainstream schools.

In terms of achievement at National level it is hard to imagine a greater level of success for an NGO project.

Twenty-three CRANE staff members also discussed the Sustainability Scorecard indicators and scored them individually. The results are aggregated in Table 60. It looks as if the staff share the views of the Mentors that engagement between schools, students and parents has improved (overall figures in **bold**). This doesn't match the findings in the Household Survey on the frequency of contacts between schools and PCGs. The overall scores on Child Protection seem too variable to interpret although they are clearly positive – consider that if all the staff members had added one point to their scores the aggregate would have risen would have risen by 23 points. That child protection at community level scores have risen so much means that there must be greater confidence among CRANE staff in this exercise and the Mentors. The lower increases in the System indicators are partly because a number of staff abstained from scoring. It would be very useful to hear from the participants – it's always true when using scalar indicators that the conversation provides more learning than the scores.

[71]Kirsten Kristensen, Martin Omagor-Loican, Negris Onen, Daniel Okot (2006) *Opportunities for inclusion? The education of learners with special educational needs and disabilities in special schools in Uganda,* British Journal of Special Education, Volume33, Issue3, pp139-147.

Table 60 - Sustainability Scorecard indicators by staff				
		2017	2019	Differences
Community Family	1 Material support to girls' education	25	61	36
	2 Involvement in school	22	57	35
	3 Moral support for education	20	55	35
	4 Child protection	28	62	34
Community Leaders	5 Engagement with schools	27	71	**44**
	6 Moral support to girls' education	25	63	38
	7 Child protection	30	70	40
School	8 Engagement with parents	26	71	**45**
	9 Engagement with students	29	76	**47**
	10 Teaching methods	19	60	41
	11 Child protection	26	61	35
	12 School administration	25	49	24
	13 Management of teachers	23	57	34
	14 Inclusion of CWD	14	50	36
System	15 Local	17	49	32
	16 District	17	49	32
	17 National	14	50	36

The Girls Themselves

The Open Qual interviews with girls and their PCGs [40-100] provided a great deal of positive discussion about confidence and ambition of the girls and of their carers for their girls. One part of the interviews invited the girls to directly address their level of confidence on a Spectrum Line and to compare their feelings with how they felt a year ago. The conversations that followed revealed important positive changes. An important story from several parents was that they had decided to send other children to school after seeing the effects on a daughter who had been helped back to school by the project. It's an interesting story for two reasons – first, it confirms part of the project theory of change (girls helped back to school do well) and, second, it demonstrates that the theory of change does not need to follow a model in which attitude change leads to behaviour change. In these cases, the carer has accepted that a girl (usually a daughter) can, for example, go to a CLC. After she leaves the CLC the girl continues in school and the carer changes her attitude to the education of other children.

The girls talked a lot about improved relationships with their parents. "They listen to us more and they talk to us about our future. They have encouraged us to get skills and not to look down on some types of work". The girls also say that their parents have taken up work and started new enterprises. They are not so "jumpy" about boys and don't think about boys all the time as some girls do and as we used to do. Their friendships have changed too and they are closer now to those who encourage and help them and also have genuine aspirations not those who are talking behind you.

The main point that comes out of the Open Qual work is a much greater sense of confidence and ambition and that these changes are likely to improve the sustainability of the activities and the impacts.

The changes in schools reported by the Learning Support Teachers and Headteachers are mentioned in section 6.3 on Teaching Quality also should be looked at in terms of their sustainability. The comments contain several themes which seem to inter-relate and reinforce each other. The use of more positive discipline methods and less physical abuse is supported by making teaching more interactive and by more engagement with the children in activities like debates, mock election, skits in assembly, talking about books they've read, acting out characters from books, writing projects[72], creating room for expression and more music, dance, drama and sport. These combine with other classroom teaching methods to improve the girls' confidence. Confidence was an issue and LSTs [5-9] mentioned girls who previously would not look you in the eye or would be too scared to answer when they know the answer or would be ashamed to speak because they thought their English wasn't good enough. This set of changes leads a headteacher [4] to say that the children are more friendly now and the LSTs to say that there is much better rapport. And my favourite quote [7] "*there is less fear of teachers and mathematics*".

It is not clear that these changes lead to better scores in learning tests but it is clear that it leads to better experiences of school and higher levels of confidence which conforms to the project theory of change and the overall purpose and the sustainability of those impacts.

Other issues of sustainability

The EET has repeatedly engaged the project staff with other issues of sustainability which are not addressed in the logical framework or the Sustainability Scorecard. These mostly concern the mobile library, IT services and transport functions. In all three cases the vehicles being used cannot be maintained and run without the project funding and the functions cannot be considered sustainable. The project staff are aware of the difficulties and have discussed them with the Fund Manager.

The library function involves the delivery of books and other reading material but also support and encouragement to children in their reading and to staff of mainstream schools in helping children with their reading. The same is true of the IT bus which means that the replacement of the buses with other means of transporting books and computers would only be a partial substitute. The issue of providing transport for children with disabilities is also not entirely straightforward. Nevertheless, the project should be experimenting with alternatives: sharing taxis; moving books by motorbike; local pools of books and computers shared by neighbouring schools; teachers taking different roles in supporting literacy and technology; … in fact, a range of ideas could be tested at the same time in order to develop proposals for maintaining or replacing what the project does with the large vehicles that could be supported in the future by the schools or the education services.

Infrastructure

We have not examined questions relating to the sustainable use and impact of project interventions in school buildings. We expect the issue to emerge towards the end of the project funding period and that there would be very few issues in the first few years. The EET will explore this area after the immediate pressure of the Midline Evaluation has been lifted.

[72] All these examples are verbatim reports from interviews.

The following sub-section and Table 25 should be completed by the project.

Set reasonable expectations: At each of the three levels of sustainability, what changes still need to take place to ensure that attitudes, behaviours or approaches are established which provide for ongoing learning and successful transition for future cohorts of girls and boys? Who are the stakeholders involved in these changes? What are the factors that help or hinder changes? Refer to your sustainability plan, Theory of Change and logframe. Be brief in the table and provide narrative analysis below the table that refers back to the mixed-methods analysis under 1).

(Table 25) Changes needed for sustainability

	Community	**School**	**System**
Change: what change should happen by the end of the implementation period?	• Families will always provide material support for all children in education • Primary caregivers will always go to school for information about their girl's performance • Families will always provide support at home for children to be learning • Families will always provide protection from abuse for their children • Community Leaders will always engage with schools to support education • Community Leaders will always promote education for all children • Community Leaders will always provide protection from abuse and support for children who have been abused	• Schools will always engage with parents about the education of children • Teachers will always engage with children to promote further learning • Teachers will always use creative, learner-centred methods • Schools will always provide protection from abuse for their children • Schools will always meet the minimum standards set by MoES • Head teachers will always support their teachers to provide quality education • Schools will always provide inclusive education for CWD	• Local community leaders will always work together to build a collective response between communities, government and other institutions that achieves better quality education • District Leaders will always lead or manage systems that promote successful elements of the project strategy for girls in private and government schools • National Government will replicate successful elements of the project strategy into their policies and practice
Activities: What activities are aimed at this change?	• School Enterprises • Family Economic Empowerment and Savings Group monitoring • Family Learning Days • Positive parenting training • Community Child Protection Committee training and support • Support of families and girls through community mentors	• Sports, Music, Dance, Drama, ICT skills development in school timetable and extra-curricular timetable • Building a love for reading through fun activities and guided reading • Building a love of maths through guided interactive kinaesthetic activities • Training of payrolled and some other mainstream teachers in creative pedagogy and safeguarding • Development of community child protection committees • Establishment of robust operational	• Child Safeguarding support for girls who have been abused • Sharing project successes with District Education Officers and MoES • Training of preservice teachers to demonstrate alternative methods of training and practice for teaching

		policies and systems in school for Safeguarding, Risk, Governance, Financial Management, Planning, people care • Training support for schools and teachers to deliver inclusive education for GWD and child mothers	
Stakeholders: Who are the relevant stakeholders?	• Parents • Local community leaders • Children • Mentors	• Head teachers • Teachers • Community groups	• MoES • JLOS – Police; judiciary; ODPP • DEOs
Factors: what factors are hindering or helping achieve changes? Think of people, systems, social norms etc.	• Poor connection to viable markets for parents • Ignorance about child abuse, prevention, reporting mechanisms • Parents who give all the house chores to the children	• Head teachers who are overstretched and poorly focussed and ill equipped for running schools. • Underqualified and/or poorly trained teachers	• No mechanisms to help different agencies work together • Limited resourcing for in-service training for government officers

Provide narrative analysis here of the points raised in the table above. Explain the change the project intends to achieve. Highlight cross-cutting activities, stakeholders and factors, but also those that relate to only one level of sustainability. Link the analysis here with that under question 1 above drawing on the scores given for each level. Link the analysis to the other Outcomes and Intermediate Outcomes.

The project wants to see in the schools in which it has worked parents, children, teachers and community leaders working together to provide safe homes, schools and communities where children feel safe, learn through creative, engaging methodologies that build skills that will empower the children with analytical skills, critical thinking skills and life skills - skills that will improve their life chances and give them the opportunity to achieve more than they had expected before they were engaged in the project.

As well as the project having a strong emphasis on providing support structures within the classroom environment that increases learning outcomes, it also puts significant emphasis on supporting girls who have been abused and preventing any abuse from happening in the future. We believe that abuse is a toxic stressor and children will not learn when living and learning in abusive environments. Therefore, the project will seek to demonstrate that freedom from abuse and fear of abuse gives higher learning outcomes and therefore better transition results. Tackling persistent issues of abuse in schools and communities requires effort for different entities to work together. As networks, Viva and CRANE have particular expertise in helping people to work together for a common purpose and we will continue to work to help different entities see how they can work together for children's issues.

Safeguarding is a crosscutting activity and engages all levels of stakeholders. There needs to be concerted efforts of everyone working together to keep children safe because when children live without violence, they will have a conducive learning environment. We believe that this is possibly one of the most important factors in promoting learning. This will require continued parental training, child empowerment, community child protection training, JLOS engagement and training, and teacher training, up to the level of MoES who need to be far more intentional about disciplining teachers who abuse children.

Teacher quality improvements is another critical priority whereby teachers are equipped with alternative discipline strategies and more creative, engaging, enquiry-based learning approaches. We need parents to be convinced by evidence that this is better than rote learning and we will do this by demonstrating improved learning outcomes in the schools that are excelling the most because of a change in teaching pedagogy and safeguarding. We will work with headteachers, DEOs, and the MoES to continue to build this philosophy and evidence base and to work out a low-cost model for rolling out such strategies across other schools.

6 KEY INTERMEDIATE OUTCOME FINDINGS

6.1 Attendance

The Baseline Report (p13, p31, p73 and Footnote 13) questions the appropriateness of using overall attendance data to assess the effectiveness of the project. The argument is that the project works with small numbers of girls in schools with large populations and the overall attendance figures cannot be influenced by the effectiveness of the work with the project girls. Changes in the attendance of project girls would not change the overall figures significantly. The FM accepted this reasoning and the methodology was changed to focus on attendance of the GEC girls. A mobile phone-based app was used to allow teachers to record the presence or absence of the girls. Spot checks were carried out to compare class registers with the actual presence of the girls.

Table 61 - Attendance IO indicators and main qualitative observations

IO	IO indicator	BL	ML Target	ML	Target achieved? (Y/N)	Target for next evaluation point	Will IO indicator be used for next evaluation point? (Y/N)
Attendance	% project girls regularly attending school, CLC, etc.	80%	83%	89%	Y	86%	Y
	PCGs say barriers reduced		20%	77%	Y	35% of respondents who say barriers have reduced	Y

Main qualitative findings

- 77% of Intervention PCGs say attending school has become easier since Baseline. C=68% p<0.02.
- 40% of PCGs who say attending school has got easier cite paying fees as being easier.
- 20% of PCGs who say attending school has got easier cite positive attitudes of girls (link to better teaching methods and to less corporal punishment which links to better school management).
- Some PCGs cite more income and savings. Some cite more flexible, patient, understanding schools.
- PCGs mention better support from family, CRANE and school.
- Head Teachers say enrolment is up because of CRANE Positive Parenting work.
- LSTs say they call Mentors or Parents when GEC girl is absent.
- Spot checks show high levels of attendance. In intervention most absences can be explained.
- In CLCs all absences can be explained.

Data from the Household Survey

The Household Survey asks the PCGs of girls who are in school if their girl attended on most days that the school was open during the previous term. At Baseline roughly 90% said their girl had been in school and 10% said that they had not.

Table 62 - Girls who attended school "on most days"

	Baseline		Midline	
	Attended on most days		Attended on most days	
Control	89%	10%	92%	8%
Treatment	85%	12%	89%	11%

In both surveys, slightly more Control girls attended and the overall attendance appears to have increased between Baseline and Midline – but none of these differences is significant. There are no differences between attendance at rural or urban schools[73].

Of those who said their girl had not been in school on most days, the reasons given at Baseline were inability to pay fees (75%) and illness (25%). These figures are not different to the Midline results where 61% mentioned fees and 30% mentioned illness. There are more "other" reasons mentioned at Midline including transport (4%), lack of interest on the part of the girl (3%) and rain (1%).

There are no differences between control and intervention in the frequencies of the reasons given for not attending.

These figures suggest that overall 5% of girls have absences because of lack of fees and 3% have absences because of sickness according to the responses given by their carers. The consistency of results between Baseline and Midline supports the credibility of these observations.

Data from Spot Checks

A set of spot checks were carried out in March 2019 in primary and secondary schools and also in CLCs and Vocational Training Centres.

Spot checks were repeated in August 2019. The work was more restricted and focused on those schools and other institutions which had relatively good registers during the earlier spot checks. This approach was based on the idea that one does not get good information on attendance where the registers are not up-to-date and completed on a daily basis.

The overall picture gained from the spot checks is that attendance by GEC girls is between 80 and 90%. See Tables 63 and 64.

Table 63 - Spot Check data for First Term 2019

	Physical Presence			
	Yes	No	Total	
African Hearts Junior School	29	1	30	97%
Kasengejje P/S	27	3	30	90%
Kasengejje SS	21	9	30	70%
Kisimbiri P/S	29	5	34	85%
Christian Friends P/S	12	1	13	92%
Mwebaza High School	16	5	21	76%
Goshem Christian P/S	30	8	38	79%
Kirema P/S	17	3	20	85%
Fort Jesus High School	10	17	27	37%
New Springs of Hope.	7	2	9	78%
Holy Family	17	11	28	61%
ROJ SS	28	1	29	97%
Munkabira P/S	21	3	24	88%
St. Charles P/S Bukerere	29	1	30	97%

[73] Chi-squared - p=0.86

St. Kizito P/S	16	4	20	80%
Divine Hope P/S	17	12	29	59%
Lugazi Model P/S	22	8	30	73%
St. Andrews S.S	13	6	19	68%
Bright Trust P/S	24	5	29	83%
Lugazi Community P/S	24	5	29	83%
Bat Valley P/S	25	3	28	89%
Mengo P/S	28	2	30	93%
Bugabo Lake View P/S	24	5	29	83%
St. Mark Kikandwa	22	8	30	73%
Central College Kabimbiri	24	5	29	83%
Old Kampala P/S	21	6	27	78%
Namasumbi C/U P/S	9	2	11	82%
House of Joy P/S	30	0	30	100%
Hillside Junior School	24	6	30	80%
Kasubi Family P/S	7	0	7	100%
Nakivubo Blue P/S	27	3	30	90%
Our Lady of Fatima SS	7	1	8	88%
St. Paul C/U P/S - Kyebando	27	3	30	90%
Good Samaritan P/S	25	5	30	83%
Kampala School for the Physically Handicapped	9	1	10	90%
Mulago School for the deaf	14	2	16	88%
Earnest P/S	9	12	21	43%
Kitebi P/S	26	4	30	87%
Paul Mukasa S.S	17	13	30	57%
Totals	**784**	**191**	**975**	**80%**

The timing of the Spot Checks in the second term of 2019 coincided with national exams in secondary schools and the data collected were not meaningful. The state of the registered prevented any possible calculations of attendance rates among GEC girls.

The data for a small sample of primary schools and for CLCs and TVET institutions are shown in the following tables.

Mainstream schools

The weighted average of observations for mainstream schools in March was 80%. However, follow up enquiries revealed a number of areas where underestimates were being made. The repeat visits in August to six primary schools returned a weighted average of 91%. Even in these schools there were problems with registers not up-to-date and only partially filled in. Where the class teacher was not present the register could not be seen at all.

Table 64 - Spot checks in Primary schools – Term 2 2019			
	Absent	Present	
Primary school 1	2	34	
2	2	35	

3	3	10	
4	0	15	
5	1	23	
6	5	24	
Totals	13	141	91%

As at Baseline, there is considerable variation in attendance rates when disaggregated by grade. It is hard to put any meaningful interpretation on this. At Baseline, the EET tried to find differences that were significant and tested the view that attendance was lower in early secondary grades but this could not be substantiated. The same applies to the Midline data for GEC girls.

CLCs

The weighted average for attendance in CLCs from the March observations is 67% and this was found to include a number of important errors. In one CLC register the girls in the vocational centre were recorded on the same page as the girls in the CLC although they were not working in the same space. Girls who had graduated to the mainstream school associated with the CLC were still recorded in the CLC register. Some girls who had dropped out were also still in the CLC register[74]. The data from the CLCs where no such errors were identified gave a weighted average of 90%. In the repeat visits in August to four CLCs which were thought to have better data, the weighted mean attendance on the day of the spot check was 98%. This high figure may not be as important as the observation that the staff of the CLC were aware of the situation and knew who was absent and why[75].

Table 65 - Spot checks in CLCs – Term 2 2019			
	Absent	Present	
CLC 1	0	12	
2	0	6	
3	0	13	
4	1	17	
Totals	1	48	98%

There are, however, genuine concerns over attendance at CLCs where most of the barriers to attendance would appear to have been removed. Since the girls receive free meals and do not need to pay and receive high-quality teaching in good surroundings with an excellent student to teacher ratio it would be easy to assume that attendance would be very high. The girls, though, are those who usually have not had a habit of going to school and routine attendance may require some support. It is also true that the CLC girls are tending to be older now than in GEC 1 which is partly a result of so many girls having already been through CLCs that bringing in sisters of those identified in GEC1 involves looking at older girls. These older girls are less likely value the literacy or numeracy offered by the CLC and may have other activities to go to compared with younger girls.

Vocational Training Centres

The spot checks in March returned an average attendance figure of 67% for the training centres. As with the CLCs, the data were shown to be influenced by cases where training was taking place away from the centre and the register had not be kept up-to-date. Registers are maintained with even less rigour in the training centres than in mainstream schools. Only one of the eight visited in March had an up-to-date register that was completed on the day of the visit.

[74] These errors were resolved because the CLC staff knew the reality of the situation which could not be discovered from the registers themselves.
[75] In fact, the few absentees were known to be unwell on that day.

Four vocational centres were visited again in August and the weighted average on this occasion was 93%. There is less pressure on the vocation training centres to maintain accurate records of attendance and there are numerous occasions when the trainees will be learning or practising a trade away from the centre. Two of the centres visited in August could not provide registers and a third showed a register that contained an accurate list of names but the pages were not dated for daily recording of attendance.

Table 66 - Spot checks in Vocational Training institutions – Term 2 2019			
	Absent	Present	
TVET 1	0	23	
2	2	14	
3	1	11	
4	1	7	
Totals	4	55	93%

Data from the attendance app

The CRANE project has developed an app that works on a mobile phone running an android operating system. The app presents the teacher with a list of GEC girls' names and allows the teacher to mark them as present or absent. CRANE staff expect the teachers to upload the data each Friday, either by joining a Wi-Fi network or by sending the data over the telephone network.

The new tool is an appropriate response to the change in approach in assessing attendance. The tool offers an easy way to check on presence or absence of GEC girls and, if used correctly, it should provide adequate quantities of reliable data to make assessments of attendance and deliver information on individual girls sufficiently frequently to allow follow-up on issues of non-attendance.

In its first full term of activity the tool was used to make over 68,500 observations of 1,186 girls' attendance in 33 different institutions. The data have been disaggregated by institution and by grade. The figures are variable and there are obviously a number of cases where the observations are unreliable. Data, for example, from locations where spot checks revealed discrepancies between the register and the girls' attendance; failure to take into account that the girls were attending at a different location and similar weaknesses are included.

Table 67 - Attendance data from phone-based app					
		Range of attendance % by school		Range of sample size (girl/day observations)	
	Average	Highest	Lowest	Highest n	Lowest n
Seven CLCs	67%	97%	52%	1903	132
4 schools P1	84%	100%	78%	495	56
9 schools P2	74%	98%	51%	675	18
18 schools P3	72%	100%	32%	1323	23
22 schools P4	82%	100%	68%	1542	41
23 schools P5	86%	100%	41%	2680	32
22 schools P6	81%	100%	44%	1678	26
21 schools P7	91%	100%	52%	3309	46
4 schools S1	56%	84%	48%	1050	104
5 schools S2	83%	97%	81%	895	186
5 schools S3	90%	99%	84%	964	104
4 schools S4	90%	92%	73%	699	52
2 schools S5	83%	100%	82%	520	40

Table 67 shows the high overall rates of attendance of GEC girls and the wide variations in averages calculated for individual schools. Note also that there are huge variations in the numbers of observations being submitted from different schools. Schools which only manage to submit 50 observations in a term (this is equivalent to reporting on ten girls over five days) are probably not getting on well with the new tool.

Very low averages in some schools do not depress the overall attendance rate for the grade where the number of observations is low. The low figures produced are probably due to some of the same difficulties reported in the Spot Checks and other difficulties with the tool, the telephone or accessing the internet.

Difficulties with attendance data

There appears to be no incentive or reward for teachers to maintain an accurate and up-to-date register or to exploit or share the attendance data that might be collected. This is a problem that is largely beyond the control of the project.

Seven of the forty-nine schools visited had up-to-date registers that had been completed on the day of the visit. The situation may be worse in primary schools than in secondary schools. Of the thirteen schools visited during the spot checks in March that did not have up-to-date class registers, 10 were primary schools.

There are currently a large number of weaknesses in the data on attendance in addition to the failures to maintain accurate registers, so that all the names in the register are those of girls who would be expected to attend, and to take the register on a daily or twice-daily basis.

The reporting from the phone-based app is subject to other weaknesses. The most common is that some dates are skipped and no data is provided; there are also cases of double-reporting on the same days and reporting on impossible days like weekends when the school is closed and dates into the future. The teachers don't always report and complain that there is no internet connection or no data allowance or simply that it is not their responsibility[76] to supply data in this form.

The data for each girl is made available to M&E staff as an excel sheet for each girl and must therefore be manipulated to provide a report for each institution. The data need to be cleaned and interpreted as shown in the example in the next paragraph.

The data from MIFA were examined in some detail and cover the attendance of 16 girls over 18 days. The data on four of the girls appears to stop during the 18 days and it is not clear how this should be interpreted. If the girls are registered and expected to attend, that is, they are genuinely absent, the attendance rate is 78%. If the girls have abandoned the course and are no longer expected to attend then the attendance rate is 93%. This example shows the difficulty of interpreting blanks in the data – they could mean "no data" or they could mean "girl is absent".

The app has significant potential and will provide important and useful data on attendance. The major problem remains that teachers have no incentive to maintain accurate records and it will take some time and cooperation with headteachers and District education staff before the situation improves.

Interpretation of the data will always be required. So far investigations of low levels of attendance have identified weaknesses in the data or the management of registers rather than high levels of absenteeism.

[76] This is heard from mainstream teachers, not from CRANE LSTs.

It is not clear if attendance data of GEC girls can be used to assess or compare the performance of schools, CLCs or training centres. In many of these institutions the numbers of GEC girls are small and changes in percentage attendance is not a helpful metric. Although the sample sizes in grades P4 to P7 are large the girls are distributed among a greater number of schools. This issue is a result of the approach based on the GEC girls which means that the school is not the strategic unit of the project although for many activities, e.g. teaching quality, school management and child protection, results should be assessed on the basis of the institution.

6.2 School Management

Table 68 - School Management IO indicators and main qualitative observations

IO	IO indicator	BL	ML Target	ML	Target achieved? (Y/N)	Target for next evaluation point	Will IO indicator be used for next evaluation point? (Y/N)
School governance	# Schools/ Centres that demonstrating better school management as they work towards 'QIS for Schools' Accountability or Foundation Verification (3-year process)	Nominal score of zero	All trained project schools provide evidence of changes made to improve school management and are working towards verification, as per the QIS Standard	All partner schools have received training. All establishing basic systems & policies. All are aiming for external verification in 2020	Y	10 Project schools achieve 'QIS Accountability' and 20 achieve 'QIS Foundations'	Y

Main qualitative findings

- All QIS scores have increased in all components and indicators. Most component aggregate scores have increased by about 30 points. But where the project has focused more attention aggregate increases are higher – Child Protection +47; People Care +38.
- Overall distribution of QIS scores has shifted towards high end of range.
- PCGs saying that their school is *Extremely well managed* up from 15% to 25% since Baseline.
- PCGs saying that management of their school has improved = 80%.
- PCGs saying that performance of Headteacher is *Excellent* up from 24% to 36% since Baseline
- Schools having a Child Protection Policy up from 60% to 96% since Baseline (Control = 90% though no difference is expected between Intervention and Control)
- PCGs agreeing that all abuse is reported in their school T = 91% C= 85% ($p<0.05$) and overall confidence in both T and C is down since Baseline probably as a result of raised awareness.
- PCGs who agree that *Teachers cooperate well with Parents* remains above 95% - no change since Baseline.
- PCGs who agree initiatives of the PTA /SC of value to their girl above 95% - no change.

The QIS programme for School Management

Improvement to school management involves training and other support by the CRANE project. The support follows a sequence of initiatives in a programme of learning called QIS. There are six modules containing 29 aims and a scoring system which forms an automatic monitoring system. The schools make progress through the different modules and reach a level where they can qualify for Foundation Level status or the higher-level Accountability which is an internationally recognised quality measurement.

Three of the modules (child protection, financial management and people care) make up the only area of support offered to Control schools and is a potential source of "contamination". The design has been approved by the Fund Manager and we assume that the risks of compromising results of the project have been assessed as unimportant. The training events are held separately for Intervention and Control schools. It seems clear that the EET should not expect to find major differences in improvement in some areas of school management where both Intervention and Control schools have received similar inputs.

Each school will receive four rounds of training before they are tested and if they reach an aggregate level of 2 (out of 4) for the different aims they are considered ready to be face the Foundation level tests or Accountability Level verification. In a change to the initial programme, the training has become residential which addresses the difficulties of getting the commitment and attention of school staff, especially headteachers, when the work is done at the school site. Observations like this are encouraging as they suggest that the project staff are alert to problems and inventive in finding solutions.

Table 69 shows an example of an area of work in the QIS programme and the scoring of the components. This is an area where the project has put a significant amount of work and the uplift in scores for the 52 schools is positive – each school adding on average nearly a point to its scores over the period that is covered.

Table 69 - Example of scoring in the QIS programme

Foundation for People Care									
Aim 1 We work together as a team to achieve our goals. We respect each other, listen to each other and support each other.		**Aim 2** Every person's work is valued, whether paid staff or volunteers. We regularly take time to show people that their work is appreciated.		**Aim 3** Every worker knows what he or she is expected to do and what his or her responsibilities are. Every worker knows who to ask for advice and feedback about his or her work.		**Aim 4** Workers get some training to help them do their work as well as they can.		**Aim 5** Workers are kept safe and healthy at work.	
2017	**2019**	**2017**	**2019**	**2017**	**2019**	**2017**	**2019**	**2017**	**2019**
69	109	79	109	77	113	69	112	63	102
	40		30		36		43		39

The progress in the different areas of the Foundation level QIS is shown in Table 70. The full table of data allow the CRANE staff to monitor change in the different schools for each component of the programme. The table shows the relatively large amount of change that has occurred in child protection which corroborates the observations made by the Mentors on child protection in schools.

Table 70 - Overview of progress in QIS Programme

		2017	2019	Change	Average change
People Care	**Aim 1**	69	109	40	

	Aim 2	79	109	30	
	Aim 3	77	113	36	38
	Aim 4	69	112	43	
	Aim 5	63	102	39	
Governance	Aim 1	88	114	26	
	Aim 2	91	116	25	
	Aim 3	97	117	20	24
	Aim 4	83	109	26	
	Aim 5	88	112	24	
Child Protection	Aim 1	80	124	44	
	Aim 2	70	117	47	
	Aim 3	71	116	45	47
	Aim 4	60	106	46	
	Aim 5	68	118	50	
	Aim 6	63	110	47	
Financial Accountability	Aim 1	67	104	37	
	Aim 2	78	113	35	
	Aim 3	74	104	30	33
	Aim 4	65	96	31	
	Aim 5	52	86	34	
Project Planning and Design	Aim 1	66	99	33	
	Aim 2	62	90	28	
	Aim 3	66	93	27	30
	Aim 4	70	100	30	
	Aim 5	66	99	33	
Child Well Being	Aim 1	72	102	30	
	Aim 2	77	103	26	32
	Aim 3	71	112	41	

There has been less progress in project planning because it has not been a focus of the training done with schools in the last year. However, the low levels of change observed in governance probably have a different explanation which is how difficult it is to change management practice within schools. This is partly due to the human condition of resistance to change which is made more difficult for people who are very busy. It is also partly due to the uncertain power structures within schools where headteachers, directors, members of the school board, the chair of the board and in some cases other bodies like a church or a foundation committee all assume responsibilities for the direction and running of the school. This assessment of bringing about change in school management will be shown to be accurate if this component of the QIS programme is still lagging behind the others in a few years' time.

Overall, the QIS programme appears to be making significant progress and the monitoring makes the progress look realistic. Figures 9 and 10 show the bulk of marks have moved up during the GEC-T project. It may be some time before the improvements in governance make an impact on the lives of the children in the schools but the changes in child protection are likely to be having an effect already.

Figure 9 - Most QIS scores in 2017 near the mid-range

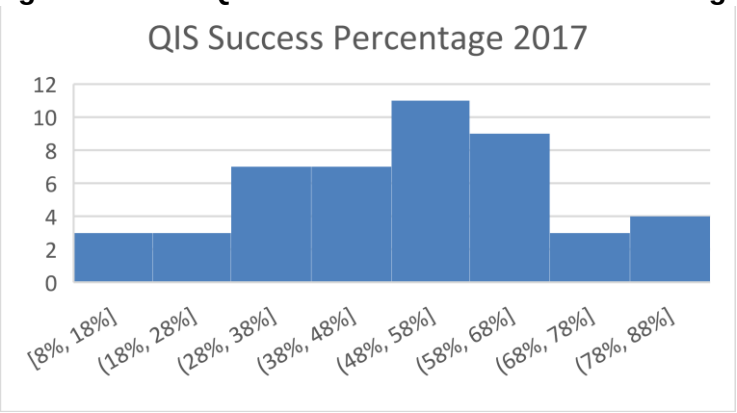
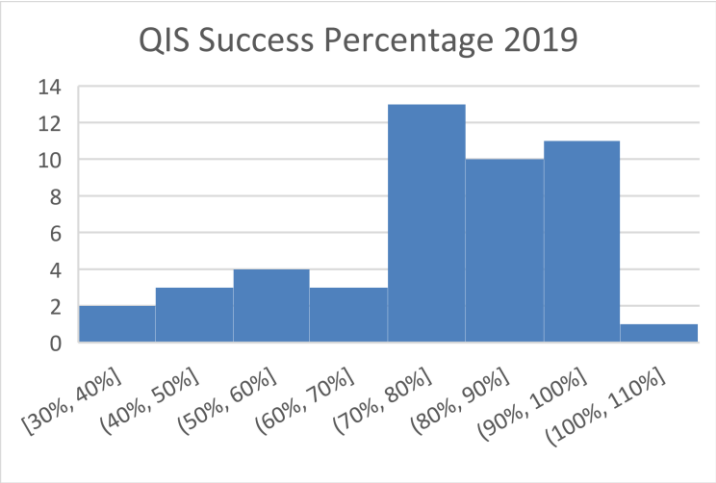


Figure 10 - Most QIS scores in 2019 above 70%



The Household Survey on School Management

The Household Survey asks PCGs how well they think the school is managed. At Baseline, we were concerned that the question design did not leave options for those who were not very pleased but could not agree on the very negative "Not at all well managed" option.

Table 71 - How well is the school managed? Baseline			
	Extremely well managed	Well managed	Not at all well managed
Control	10%	68%	3%
Intervention	15%	62%	3%

We were also positive that the distribution of responses gave room for improvement from *Well managed* to *Extremely well managed*.

Table 72 - How well is school managed? Midline				
	Extremely well managed	Well managed	Fairly well managed	Not at all well managed
Control	25%	59%	15%	1%

Intervention	25%	61%	13%	1%

Both these reflections at Baseline were shown to be true. First the option of Fairly well managed attracted a significant number of responses and, second, the proportion in the highest category of satisfaction grew. One can see by eye that there are no significant differences between Intervention and Control responses but for the record the chi-squared test returns a p-value of 0.89.

In the Baseline report we presented a table which compared learning test results for those schools said to be *Extremely well managed* with those from *Fairly well* managed schools which showed no important differences. In its place we present a table of learning test results disaggregated by the PCGs' rating of the performance of the Headteacher.

Table 73 - Learning Test Results by performance of Headteacher					
	Mean Literacy Marks	Change since Baseline	Mean Numeracy Marks	Change since Baseline	n
Excellent	36.3	13.2	33.5	9.9	261
Good	40.2	14.0	37.6	11.3	414
Fair	36.7	13.4	34.3	8.6	54
Poor	29.5	13.0	24.4	6.6	6

The results are seductive and show an increase in learning test performance as the ratings improve from Poor to Good. Unfortunately, the Poor ratings are based on only six observations and the results for Excellent Headteachers are lower than for Good. We are left with modest increases from Fair to Good

The question regarding PCGs' views on how school management has changed over the last year yielded almost exactly the same responses as at Baseline. Almost all respondents say that the management has improved. Almost nobody thinks school management has got worse.

Table 74 - How has school management changed in the last year				
	Improved	Stayed the same	Got worse	Got much worse
Control	80%	18%	1%	0
Intervention	81%	18%	<1%	0

Parents and carers were also very positive about the performance of the headteacher in their management of the school. There were no differences between the Intervention and the Control school responses either at Baseline or Midline and the data have been combined in the table below.

Table 75 - How would you rate the performance of the headteacher? Control and Intervention combined.				
	Excellent	Good	Fair	Poor
Baseline	24%	68%	6%	2%
Midline	36%	56%	7%	1%

In the Baseline report we wondered if we would see responses move from *Good* to *Excellent*[77] and that seems to have occurred already to some extent. The difference between Baseline and Midline is highly significant ($p<0.001$).

If questionnaire surveys delivered data that could be taken at face value, the project could claim to have improved school management. As with many other issues examined in the Household Survey, it is hard to know how to interpret the responses of the participants. The largest confounding factor is probably the way that the project work raises the awareness of an issue, perhaps focusing people's attention on things that they had not considered in detail before.

At Baseline, just over 60% of respondents said that their school had a Child Protection Policy. At Midline the proportion has risen to 96% in Intervention areas and 91% in Control areas. The difference at Midline is significant ($p=0.01$).

Table 76 - Does your school have a Child Protection policy?		
	Yes	No
Control	91%	9%
Intervention	96%	4%

The project has been helping schools to develop a Child Protection Policy and has worked with teachers and headteachers on child protection issues. The project could rightly claim some responsibility for the increase in child protection policies. Only a cynic would say that the difference was due to the parents and carers becoming more alert to what the "correct" answer to the question might be.

Primary Care Givers are asked if girls are safe when they are in school and offered a range of options from always safe to never safe. At Baseline, the 78% of respondents chose always safe and 18% chose usually safe. Only 6% thought girls were rarely or never safe.

Table 77 - Are girls safe in school?				
	Always safe	Usually safe	Rarely safe	Never safe
Control	80%	14%	5%	1%
Intervention	79%	17%	2%	1%

The situation has barely changed between Baseline and Midline – there is no difference between Control and Intervention in either evaluation event and no difference between Baseline and Midline[78]. The persistent small percentage who feel that girls are not safe at school remains a concern. The Ministry of Gender report[79] and Devries *et al*[80] suggest that girls face greater safety issues on their journeys to and from school.

Most cases of abuse are reported

An interesting test question on child protection issues asks PCGs if they think all cases of abuse at school are reported. At Baseline, there was a large agreement with the statement and only 7% of respondents disagreed. There was no difference between the responses in Control and Intervention areas.

[77] Baseline Report, p75.

[78] Chi-squared tests may be unreliable where expected values are very low, as in this case with the Rarely and Never cases. Combining cases and testing positive against negative responses also does not produce a significant difference.

[79] Ministry of Gender, Labour and Social Development. *Violence against Children in Uganda: Findings from a National Survey, 2015.* Kampala, Uganda: UNICEF, 2015

[80] op cit

Table 78 - Almost all cases of abuse are reported				
	Strongly agree	Agree	Disagree	Disagree strongly
Baseline (C&T)	37%	57%	6%	1%
Midline Control	35%	49%	11%	4%
Midline Intervention	45%	47%	6%	3%

Two changes appear to have occurred according to the responses at Midline. First, the Control and Intervention responses appear to have moved apart with those from Intervention areas becoming more confident that abuse is fully reported. The difference is significant (p=0.049) but given the small values in some Expected cases the result cannot be relied upon. The difference seems to be driven by the higher numbers saying that they agree strongly with the statement.

Second the overall level of confidence that abuse is being fully reported has declined. The difference between Baseline and Midline is significant[81]. This result was anticipated in the Baseline report (see page 78) on the basis that greater awareness of issues of abuse and changes in understanding of what constitutes abuse might lead parents to question the level of reporting. The two changes appear to be contradictory – on the one hand the parents in Intervention areas are more confident at Midline of more complete reporting and on the other the overall picture is of less confidence than at Baseline.

This is an area where qualitative interviews are necessary to complete our understanding of what is going on. Are people changing their ideas about abuse or are they changing the way they answer the survey? Our interviews with the project Mentors suggest that there are changes of opinion among project beneficiaries like the PCGs who took part in the Midline Surveys and that they are more aware of the prevalence of issues of abuse[82] and that they are aware of the work being done by the project to raise the issues and address them.

The PCG were asked if they thought that initiatives by the school committees or Parent Teacher Associations were of value to their girls at the school. The responses were overwhelming positive, as they had been at Baseline, with only 9 responses demurring from the positive option. The question is probably not worth repeating in any future Household Surveys.

Table 79 - Teachers cooperate very closely with parents on school matters.				
	Strongly agree	Agree	Disagree	Disagree strongly
All observations	54%	44%	2%	1%

The results of the question asking parents to agree or disagree with a statement about cooperation between teachers and parents were hopelessly positive at Baseline and the same result was observed at Midline. As remarked in the Baseline report (p79), the findings are not very helpful and it may simply be that they are the result of the Acquiescence Effect or respondents may see through the question and give the answer they assume the researcher wants to hear. At this stage, it is perhaps possible to take comfort from the fact that the responses have not become negative but overall it is probably not useful to continue to include this question in future surveys.

[81] The differences are highly significant (p<0.001) regardless of whether the data for Control and Intervention are examined separately or combined. When the data are combined there are no very small Expected values.
[82] We do not want to lean too heavily on a single report but the Devries et al paper (op.cit.) suggests that if verbal abuse is included in assessments, virtually all children (the figures are over 90%) experience some abuse.

Comparisons between government and private schools show very little difference in almost all the variables examined in this chapter. There are no differences between what PCGs say about government and private schools on: how teachers cooperate with parents; how well managed the school is; the performance of the head teacher and absenteeism of teachers.

There is, however, a highly significant difference in the frequency with which girls report having seen a child receiving corporal punishment (caning)[83]. The rates are 64% of girls in government schools and 51% of girls in private schools (p=0.001).

6.3 Quality of Teaching

Table 80 – Quality of Teaching IO indicators and main qualitative observations

IO	IO indicator	BL	ML Target	ML	Target achieved? (Y/N)	Target for next evaluation point	Will IO indicator be used for next evaluation point? (Y/N)
Teaching Quality	Proportion of project teachers who consistently demonstrate the 10 essential criteria in the project lesson observation tool	Not assessed	60% do 5 or more of the 10 essential criteria	48% Much qual evidence of improved quality. Better LT results where less "chalk and talk". Sig diffs T vs C.	N	70% do 5 or more of the 10 essential criteria	Y
	Project teachers can identify skills they are learning that make them a better teacher	Not assessed	Growing evidence that teachers can identify skills they are learning through the project that make them a better teacher	Video reports. LST interviews. T and HT interviews	Y	Growing evidence that teachers can identify skills they are learning through the project that make them a better teacher	Y

Main qualitative findings

- Girls who agree that the "*main way of teaching is for teachers to write on the board and students copy in their exercise books"* C 68% and T 56% (p=0.004)
- PCGs on same question also different C=79% and T= 69% (p=0.06). Girls probably know better than PCGs!
- PCGs say quality of teaching is *very good* up from 19% to 28% since Baseline.
- PCGs who agree *School environment is hard but there's nothing you can do* down (32% to 28%) since Baseline (not significant).
- PCGs who disagree that *Most teachers think there are some subjects that girls can't do as well as boys* up from 33% to 44% since Baseline.

[83] The interpretation of this question

- Girls who disagree with *Most teachers think there are some subjects that girls can't do as well as boys* C=T= 60%
- 70% of lessons observed meet 4 criteria in Lesson Observation tool. (Not necessarily "essential" criteria).
- 60% of lessons observed meet 7 criteria in Lesson Observation tool.
- Open Qual girls say school has improved.
- LSTs and Teachers say teaching methods have improved and mention 16 different techniques that have been used.
- HTs say schools are happier and relationships between teachers and pupils have improved. Partly due to reduction/elimination of corporal punishment. EET here seeing better sense of personal safety as a component of better teaching quality.
- LSTs say pupils have less fear of teachers and mathematics.
- Open Qual parent (who is teacher) talking of improving his teaching.
- CRANE staff talk about training Centre Coordinating Tutors to train teachers.
- Head Teachers say Library truck and IT truck improve learning.
- Head Teachers say enrolment is up because of better teaching methods.
- Videos of teachers show them doing self-critical assessments of their own practice and their ideas for improving.

The CRANE project operates a wide range of activities designed to make the school experience safer, more enjoyable and more effective in helping girls to learn. The activities may be difficult to separate in terms of their impacts as they can all contribute to better attendance and better performance.

The Household Survey was modified after Baseline so that it explores the opinions of both the Primary Care Givers and the In-School Girls (ISG) on the quality of teaching being provided in the sample schools. A few questions were added on the basis of learning from an education project in Tanzania where members of the EET are involved.

Questions repeated from Baseline include asking respondents to agree or disagree with the statement, "*Most teachers think there are some subjects that girls can't do as well as boys*". In both Control and Intervention populations the numbers who agree with this statement have increased since Baseline. The difference between Intervention and Control has become significant at Midline[84].

Table 81 - Most teachers think there are some subjects that girls can't do as well as boys

		Strongly agree	Agree	Neither agree nor disagree	Disagree	Disagree strongly
Baseline	Control	9%	16%	5%	19%	18%
	Intervention	12%	18%	4%	19%	14%
Midline	Control	11%	19%	10%	35%	19%
	Intervention	11%	29%	6%	24%	20%

At face value these changes imply that carers in Intervention areas are more likely to think that teachers are sexist in some of their teaching now than they did at Baseline and more than carers in a Control area. If teachers in Intervention areas have become more sexist in their views this would represent a serious failure of project activities. It is possible that carers in Intervention areas have become more sensitive to the issue and are more prone to think that teachers are likely to be sexist.

[84] Chi-squared – p=0.02

The girls were asked the same question on teachers' views on competencies of boys and girls and both Intervention and Control girls answered in the same way with 40% agreeing with the statement and 60% disagreeing – that is, a slight majority saying that they don't perceive teachers as sexist in this area.

The second question asked at both Baseline and Midline invited participants to agree or disagree with the statement, "The school environment is hard but you can't do anything about it." In work in Tanzania, we have found interviewees who disagree with this statement tend to be more progressive and positive in their responses to other questions about education. In both Control and Intervention samples the numbers disagreeing with the statement have increased. The changes are not significant[85] between surveys or between Intervention and Control.

Table 82 - The school environment is hard, but you can't do anything about it			
		Agree[86]	Disagree
Baseline	Control	32%	68%
	Intervention	32%	68%
Midline	Control	26%	74%
	Intervention	30%	70%

The issue on how well teachers work for all their students is addressed in the Household Survey where respondents are asking to agree or disagree with the statement, "*Most teachers only work with the best students and don't help those who have more difficulty understanding*." There was no difference between Intervention and Control at Baseline where a slight majority chose to disagree.

Table 83 - Most teachers only work with the best students and don't help those who have more difficulty understanding.					
		Strongly agree	Agree	Disagree	Disagree strongly
Baseline	Control	16%	15%	21%	19%
	Intervention	18%	15%	19%	20%
Midline	Control	16%	17%	39%	28%
	Intervention	20%	25%	28%	27%

At Midline, the two populations appear to have diverged with more carers in treatment areas agreeing with the statement (45%) than in Control areas (33%) - the difference is significant (p=0.02).

At face value, this appears to mean that now more carers think that teachers are not teaching for all the pupils than thought so at Baseline. In fact, it is likely that the opposite is true as teachers in Intervention schools adopt more "differentiation" in their teaching methods. This may be an example of the crime statistics problem[87] where raising awareness of a problem makes it appear worse even while it is being addressed.

The HHS also asked girls and carers about teaching methods by inviting them to agree or disagree with the statement, "*The main way of teaching is for teachers to write on the board and students copy in their exercise books*." The result from the girls shows a significant[88]

[85] p=0.27
[86] Numbers of Strongly Agree and Agree combined.
[87] See for example - https://fullfact.org/crime/q-can-we-still-trust-crime-statistics/ Statistics for a particular crime appear to get worse when the police focus on it. When the police are less active in the area, the crime statistics appear to improve.
[88] p<0.004

difference between Intervention and Control with 56% and 68% respectively agreeing with the statement.

The proportions of carers who agree are higher (Control = 79% - Intervention = 69%) and although there is a difference between the two groups the difference is not highly significant (p=0.06).

Intervention girls and their carers seem to think that more teachers in their schools use imaginative methods than their counterparts in Control areas. If the girls in Intervention areas are correct then the project might be able to take some credit for introducing teachers to more interesting and engaging teaching.

It is rather difficult to interpret these data without reference to the results of qualitative interviews. We are aware that it will seem tendentious if we say that positive results are due to project successes and negative results are due to the higher levels of awareness.

A major concern in teaching quality is absenteeism among teachers. This problem was raised in Baseline qualitative interviews and a question was introduced into the Household Survey to assess the seriousness of the issue.

The issue is described in a report by Allen *et al* (2016)[89] where absenteeism is assessed in two ways: absent from the classroom and absent from the school. The numbers of Kampala are 38.1% and 10.8% respectively. The rates in other parts of Uganda a worse or far worse than in Kampala. The UNICEF[90] report (2016) shows how absenteeism and "lateism" by teachers in primary schools correlate strongly with the same issues among pupils – if the teachers don't turn up the children don't turn up either.

Table 84 - Teachers in [girl]'s school are often absent and [girl] gets no teaching when the teacher is not present.

		Strongly agree %	Agree %	Disagree %	Disagree strongly %
Girls	Control	6	13	43	38
	Intervention	6	11	42	41
PCGs	Control	5	10	73	13
	Intervention	5	6	66	22

There are no significant differences between the frequencies of responses from Intervention and Control areas. Parents and carers seem to see the problem as being less prevalent (9%) than the girls (18%) – this difference is highly significant (p< 0.001).

The Household Survey contains a suite of questions on teaching quality which have been asked at Baseline and at Midline. Some of these questions do not seem to be adding much to our learning.

		Baseline		Midline	
Table 85 - Have you been informed about [girl]'s progress in the last 12 months?		Yes	No	Yes	No
	Control	89	11	90	10
	Treatment	88	12	89	11

[89] Reg Allen, Phil Elks, Rachel Outhred and Pierre Varly (2016) *Uganda's Assessment System: a Road-Map for Enhancing Assessment in Education*, HEART, 14 September 2016

[90] https://www.unicef.org/esa/sites/unicef.org.esa/files/2019-05/UNICEF-Uganda-2016-Absenteeism-Key-Driver-Poor-Performance-Primary-Education.pdf

The results are remarkably uniform and stable. Nine out of ten PCGs have been informed about their child's performance.

When we ask the slightly different question about how often the PCG receives information from the school about its plans we also get uniform and stable responses (see Table 62). This question fits with issues of school management rather than teaching quality but seems to follow on logically from the previous question about being informed about an individual girl's progress. If you have looked at the percentages in the table you already know that there are no significant differences between the overall figures for Control and Intervention or between Baseline and Midline. However, if you do the somewhat unorthodox move of treating only the data for monthly and termly contact at Midline, there is a significant difference (p=0.02) – PCGs whose girls are in Intervention schools say they have more frequent communication than those whose girls are in Control schools. The difference is not significant at Baseline.

		Baseline		Midline	
Table 86 - How often does the school communicate with you about its plans and activities?	Frequency of contact	Control	Intervention	Control	Intervention
	Weekly	1	2	1	1
	Monthly	15	16	11	17
	Termly	70	67	79	68
	Yearly	8	9	7	7
	Never	6	8	3	6

Another question that delivers uniform and stable response frequencies is when PCGs are asked if they are aware of changes in teaching practices at school. Three out of four PCGs say that they are aware at both Baseline and Midline and in both Control and Intervention areas. We need to rely on the qualitative interviews with parents to find out what they think the changes have been.

		Baseline		Midline	
Table 87 - Are you aware of any changes in teaching practice at [girl]'s school?		Yes	No	Yes	No
	Control	76%	24%	77	23
	Treatment	72%	28%	75	25

The responses to the question about how often a parent or carer has been to visit the school their girl attends are very confusing. At Baseline a large proportion of PCGs said that they had visited at least three times – the differences between Intervention and Control are not significant. At Midline the overall frequency of visits has dropped in both Control and Intervention areas. Intervention PCGs appear to be visiting more often than those in Control areas but this is because more of them make multiple visits while the proportion who make a single visit (in one term) has dropped. The calculations suggest the difference is significant but we have little confidence in the data and cannot interpret the results to explain the behaviour.

		Baseline		Midline	
Table 88 - How many times did you go into [girl]'s classroom or school?	Number of visits	Control	Intervention	Control	Intervention
	Never	18	17	23	27
	Once	20	15	31	18
	Twice	27	27	23	29
	3 or more	35	41	23	26

The qualitative interviews seem to suggest a different story in which PCGs say they are more in touch with the school than previously. The responses are perhaps from areas where more effort to be in touch is being made or where the school has a larger role in the community being perhaps linked to a church, a demonstration centre and having both Primary and Secondary elements. The interviews with the Learning Support Teachers (LST) also support the idea of more contact between teachers and parents. In these cases, the link is initiated by the LSTs following up on absences of the girls directly with carers or through the Mentor. The occasional call to ask after the girl may have a stronger effect on how connected to the school the carers feel.

Asked about the quality of teaching their girl receives, PCGs are very positive with tiny numbers choosing neutral or negative responses. This may be the Acquiescence Effect at work and it may be impossible for a parent to say that they send their child to a school where teaching is poor. There are no differences between the response rates in Intervention and Control areas.

		Baseline		Midline	
Table 89 - How would you describe the quality of education that [girl] receives?	Quality of teaching	Control	Intervention	Control	Intervention
	Very good	15	19	27	28
	Good	74	69	66	65
	Neither	7	5	5	4
	Poor	1	3	2	2
	Very poor	1	1	0	0

However, there are similar changes in both groups between Baseline and Midline as the level of approval gets higher with higher proportions of votes appearing in the “very good” category. These changes are significant for both Intervention and Control (p= 0.01 in both analyses).

		Baseline		Midline	
Table 90 - In the last 12 months, how do you think the quality of education that [girl] receives has changed?	Change in quality	Control	Intervention	Control	Intervention
	Got better	78	84	84	85
	The same	22	16	15	14
	Got worse	<1	<1	1	1

At Baseline, there is a difference between the views of PCGs in Control and Intervention areas with those in Intervention areas more positive about recent changes in the quality of teaching. The difference is nearly significant at p = 0.06. By Midline, the difference has disappeared and both groups of parents seem equally positive about recent improvements. Nobody says that teaching has got worse.

These two questions seem to corroborate one another in that all PCGs are saying that the quality of teaching has been improving with Intervention schools apparently starting earlier than Control schools. This though may not be the vindication that project staff might have hoped for after their efforts between Baseline and Midline.

The household survey results indicate that PCGs of girls in government schools have almost always the same opinions as those whose children are in private schools. There are no differences between their responses on the quality of teaching or how the quality has

changed in the last year. They have the same opinions on the use of copying from the board as the most common teaching method.

There is however as significant difference between their views on whether teachers think that there are some subjects that boys do better than girls. Although those who think the teachers are sexist in this way are a minority, the 44% who think so whose girls are in government schools is significantly different from the 33% where the girl is in private school (p=0.01).

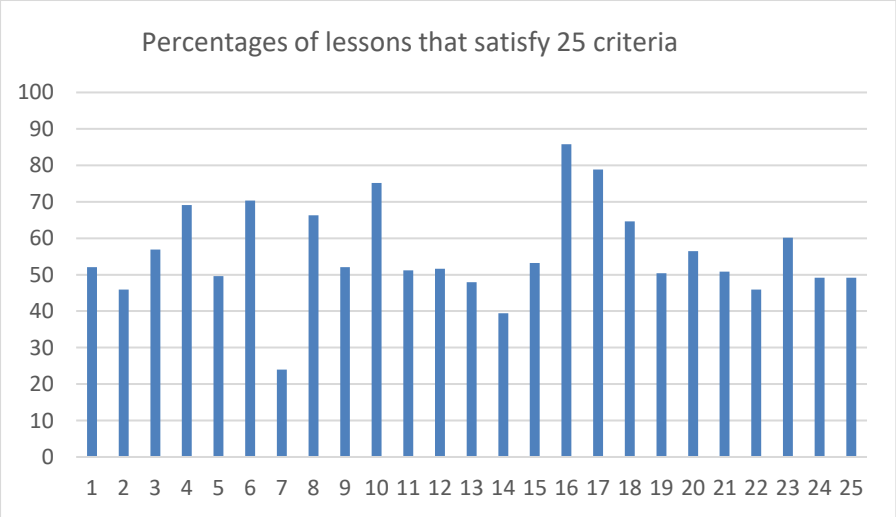
CRANE monitoring

The CRANE project staff are doing a range of activities to improve the quality of teaching. This includes work with teachers, head teachers and District education staff. There is currently a strong focus on "differentiation", that is the teaching based on the needs, interests and abilities of the learners.

The CRANE monitoring includes lesson observations where a lesson is scored according to how many of 25 criteria are satisfied by the design and running of the lesson. This forms the basis of logframe indicator (Intermediate Outcome 2). Teachers are observed on several different occasions and their personal score may improve. However, it would be possible to aggregate the observations from a large number of lesson observations and look for an overall improvement. This would allow an improvement made in any area of teaching to be recorded as positive progress and would take the pressure off individual teachers who anyway get the learning they need from the feedback from the observer.

The EET has seen a sample file from the KoBoCollect app in which observations are recorded which covers 123 lessons carried out by 103 teachers or teaching teams. The results show how often teachers satisfy the criteria. Figure 9 shows how often the lessons comply with the different criteria. The lowest level of achievement shown is in the use of ICT (criterion 7) and the highest level is establishing a positive rapport (criterion 16).

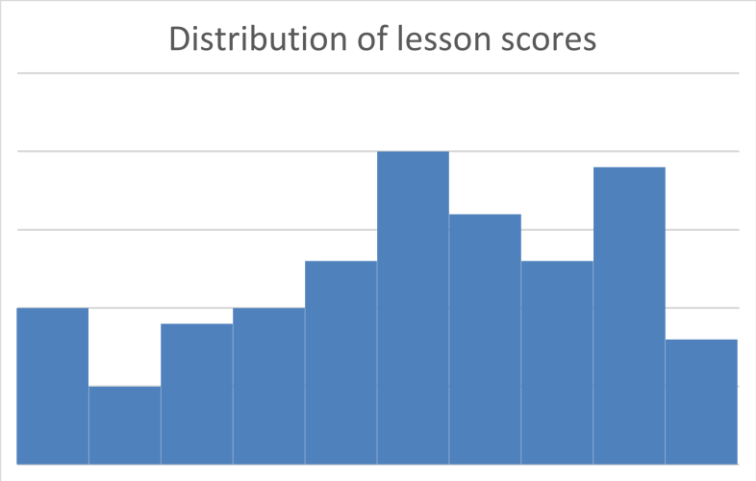
Figure 11 - How often Teaching Quality criteria are met



The same data show how often the criteria are achieved in lessons with a distribution that approximates to a Normal Distribution despite a rather high number of lessons scoring very poorly. The distribution suggests that the criteria are reasonable and that there is room for improvement that could be monitored over the next few years.

The next round of lesson observations will provide specific quantitative data as well as qualitative learning about improvements in the quality of teaching.

Figure 12 Distribution of lesson scores



It may be important to note that the monitoring is based on the lessons and teachers and it is not obvious that the data can be used to assess changes at the level of the school. It is another example of how the evolution of the project design has not led to a monitoring focus on a specific unit that unites the work of the project monitoring and the EET evaluations.

Self- awareness of teachers

Teaching quality is also assessed by the level of awareness of teachers of their own strengths and weaknesses. This is an unusual indicator and it requires skilfully facilitated interviews with teachers in what the perceive as an open uncritical environment. The interviews with teachers that we have seen on video and read in transcript certainly show a level of self-awareness and a great interest in teaching approach and methods. There tends to be a strong focus on the methods that facilitate more differentiated teaching based on the learners' competencies. This may not be the easiest area to work in given the frequent high numbers of pupils in each class.

Best teacher initiative

The promotion of improved teaching quality is supported by a best teacher competition and this has recently been modified to take the form of a prize for the most improved teacher. This represents an astute understanding of the circumstances and how the prize can be made more attractive and motivating to more teachers. The teachers were familiar with the 25 criteria being used to judge lessons and were told that their lessons would be assessed on these criteria. The assessments therefore served not only to help the teachers design and deliver their show piece lessons during the training camp but also to help them understand the criteria that are used by project staff in their lesson observations.

EET qualitative interviews

The qualitative interviews with teachers, headteachers and Learning Support Teachers are remarkably consistent in their assessments of changes in teaching practice. Headteachers are aware of higher levels of enrolment which they attribute to better teaching methods and work on positive parenting. Headteachers praise the LSTs and claim that children are learning more. Headteachers mention the computer truck as a contribution to improving learning.

LSTs all talk about differentiation (without using the word) and what a difference it makes to classwork. Better attitudes among pupils are attributed to attempts to conform to the criterion of creating a positive rapport and to attempts to avoid corporal punishment and to

use alternative positive methods of discipline. The use of several teachers is also seen as an important method for improving teaching. This "co-teaching" is also seen as a difficulty where teachers are not "compliant". The reading project is called by one LST "drop everything and read" and is seen as an amusing way to launch an initiative which is having positive results on children reading. The LSTs also mention the access to reading material from the Library bus.

The LST interviews mention a huge inventory of teaching methods that are presented as innovations of which they are proud. A short list here will provide some examples: team teaching sometimes three teachers in one class; giving out success criteria; peer assessments; self-assessments; group work; energisers; not writing in every lesson; plenaries to reflect on learning; practical activities; learning games; different instructions to different groups; allowing discussion; making children feel free to ask questions and having fun.

Less fear of teachers and mathematics

The LSTs and Headteachers do not separate their comments on child protection, support to girls with disabilities, methods of improving teaching quality, income generation initiatives and girls' confidence but seem to discuss the component issues in a continuous way. The comments are not a stream of praise as this account of positive statements might imply but there is a clear sense, shared by CRANE head-office staff, that teaching and the environment in which teaching is taking place, are improving. There is no attempt to identify which cause leads to which effect.

A teacher's lot is not a happy one

It is important to bear in mind that most teachers supported by the project work in conditions in which it is terribly difficult to be a good teacher. No one disagrees with the criteria used to assess the quality of teaching but it is very difficult, for example, to provide differentiated teaching in a classroom containing 90 children under even the best conditions. Teachers are often working in very poor conditions with very few resources with children who have very little school equipment and a huge range of capabilities. Teachers may also face headteachers and colleagues who are not supportive of more learner-centred teaching methods.

There is some corroboration of this concern of the environment in which the project is trying to introduce innovations in teaching. Teachers who attend training camps show much greater freedom and initiative than they do in their observed lessons in school. Of course, this is what one would expect but it is a useful observation, nonetheless. Also important is the fact that the teachers are making the same observation and are learning that they can successfully use different methods and don't need to experiment in their own classes to find this out.

The EET was touched by one interview in the Open Qual work which was discussing parental support to girls in education where one of the carers said that he was a teacher and felt very bad about the way his long hours at school meant that he was not supporting his daughter sufficiently.

6.4 Life Skills

Table 91 – Life Skills IO Indicators and main qualitative observations							
IO	**IO indicator**	**BL**	**ML Target**	**ML**	**Target achieved? (Y/N)**	**Target for next evaluation point**	**Will IO indicator be used for next evaluation point? (Y/N)**

Life Skills	% girls sampled who have improved at least one life skill that has been taught by the project (ICT, a Sport, Music, Dance, Art, IGA, etc.)	34%	Target 40% sampled girls gain at least one new project-taught life skill	80% of girls are in or have been in one of the 52 schools where ICT, MDD, and other extra-curricular opportunities are offered	Y	50% sampled girls gain at least two new project-taught life skill	Y More work needed to improve LS questions in Household and Girl Surveys. Open Qual work using individual and group interviews with girls needs to be repeated.
	Girls can identify skills they are learning in the project that will be useful to their lives	No assessments	Increasing evidence that girls identify skills that will be useful to them in life	*Going to school does influence what you end up doing in adult life* Strongly agree increased 5%. Open Qual evidence of decisions on activities, relationships and ambition	Y	Increasing evidence that girls will identify skills that will be useful to them in life	Y Refinements needed in information gathered for indicator - see above

Main qualitative findings

- 20% of PCGs who say attending school has got easier cite the girl's positive attitude as one reason.
- Girls who disagree strongly with *I avoid new things if they look difficult* up from 15% to 28% since baseline (and those who agree down from 35% to 24%) = more confident.
- Girls who disagree with *I get nervous speaking in front of people my own age* T= 72% C= 64% p=0.03. GEC girls are more confident.
- Girls who disagree with *I get nervous speaking in front of an adult I don't know* T= 45% C=40% p=0.01. GEC girls are more confident.
- Girls who say they make decisions about whether they go to school or not T=31% C = 20% p<0.001
- Girls who say they make decisions about what they do as a job T=80% C = 62% p<0.001
- Girls who say they make decisions about how much time they spend with friends T=62% C = 50% p<0.001

  ⇨ GEC girls have more say in decision making in all three cases than Control girls.
- Girls in Open Qual have more confidence and ambition, say they are doing new things and have better relationships with their parents. Have changed their friendship groups and have better relationships (attitudes) to boys. Their parents are doing new activities.
- Girls in survey qualitative interviews do spectrum lines that show they believe they are more confident.

Table 92 - Life is harder for girls but there is nothing you can do about it

		Strongly agree	Agree	Disagree	Disagree strongly
Baseline	Control	32%	36%	21%	4%
	Intervention	37%	31%	22%	5%
Midline	Control	37%	29%	19%	6%
	Intervention	34%	33%	17%	11%

At Baseline, the EET suggested that the balance of 68% agree and 25% disagree might provide a good starting point to detect change. However, there have been no changes in how the question has been answered. There are no significant differences in any comparisons that can be made in the table.

The question, *I avoid new things if they look difficult* is an attempt to get an assessment of self-confidence. Improved confidence may be the most important benefit that education may bring.

Table 93 - I avoid trying new things if they look difficult					
		Strongly agree	Agree	Disagree	Disagree strongly
Baseline	Control	10%	24%	48%	12%
	Intervention	11%	24%	47%	15%
Midline	Control	12%	14%	42%	27%
	Intervention	7%	17%	43%	28%

The responses are not different between Intervention and Control but the combined data at Midline are very significantly different from the data at Baseline ($p<0.001$). There has been a decrease in those who Agree and an increase in those who Disagree strongly.

Two additional questions exploring girls' sense of their own confidence are formulated about speaking in front of other people.

Table 94 - I get nervous speaking in front of people my own age					
		Strongly agree	Agree	Disagree	Disagree strongly
Midline	Control	9%	22%	36%	28%
	Intervention	8%	17%	40%	32%

Intervention and Control frequencies are not different ($p=0.19$) when the four responses are examined but when the data are aggregated as Agree vs Disagree a difference emerges which is significant ($p=0.03$). At face value, Intervention girls are more confident than Control girls.

Table 95 - I get nervous if I have to speak in front of an adult that I don't know					
		Strongly agree	Agree	Disagree	Disagree strongly
Midline	Control	27%	31%	22%	18%
	Intervention	16%	34%	26%	19%

Although the question in Table 75 is different from the preceding question and the overall difference in the balance of responses does not seem large (58:42 Control and 52:48 Intervention) the difference is significant ($p=0.01$).

Table 96 - I find it easy to tell people what I am thinking					
		Strongly agree	Agree	Disagree	Disagree strongly
Midline	Control	28%	39%	23%	10%
	Intervention	23%	44%	23%	10%

The question presented in Table 80, tries to examine self-confidence and the ability to express oneself. The responses are well split across the agree-disagree divide and Intervention and Control are virtually identical. This may be worth repeating at another evaluation event.

Table 97 - Going to school does influence what you end up doing in adult life					
		Strongly agree	Agree	Disagree	Disagree strongly
Baseline	Control	47%	40%	9%	1%
	Intervention	49%	38%	8%	1%
Midline	Control	55%	35%	4%	5%
	Intervention	54%	38%	4%	3%

We might similarly have abandoned the question on whether going to school affects what you do in adult life because it produced massive Agree vote and this seemed to tell us very little. The EET informed the CRANE project about the girls who had said that school did not influence one's adult situation and project staff followed up a number of cases[91]. The girls who were questioned said that they strongly believed that attending school affected what one did as an adult and that they had not understood the question correctly. It is possible that the same thing occurred at Midline which might explain the small numbers who responded that attending school did not have this effect.

Questions that were not particularly useful include: *I can work well in a group*; *I recognise when choices I make today can affect my life in the future; I have trusted friends I can talk to* and *I have trusted adults I can talk to*. In all these cases over 90% of the responses were on the Agree side of the options for both Intervention and Control. The same thing also happened at Baseline with questions based on *If I do well at school, I will be able to do what I want* (99% agree) and *I keep trying when others have given up* (94% agree).

These apparent failures are partly because it is difficult to find out about how confident people feel through a questionnaire survey. However, the qualitative work, especially the Open Qual interviews reveal a very positive story about GEC girls' levels of confidence and ambition. See OQ work reported under "The Girls Themselves" in Chapter 5 on the Sustainability Outcome.

Girl on Decision Making.

Under the chapter on the Learning Outcome we reported on a question in the HHS on who decides when a girl should leave school. There were three other questions on decision making that seem more appropriately treated here under Life Skills. There are also significant sections under the Sustainability Outcome that discuss other aspects of Life Skills in particular, the evidence of increasing confidence among GEC girls. The word confidence appears often in the project statements and is an aim as well as a means.

Table 98 - Who decides if a girl should go to school or start a training ?				
		Girl decides	Joint decision	Adult decides
PCG	Control	20%	17%	63%
	Intervention	31%	20%	48%

The second question about decision-making overlaps with the first and delivers the similar results but the difference between Control and Intervention is even more marked and is very significant (p=0.0002).

Table 99 - Who decides a girl's type of job or career?				
		Girl decides	Joint decision	Adult decides
PCG	Control	68%	8%	24%
	Intervention	80%	9%	11%

[91] See Baseline Report, p83.

The difference between control and intervention becomes even more marked in the decision making over the work or career the girls choose. ($p<0.0002$). The percentages follow the same pattern but the higher proportion of Intervention girls who say they make the decisions themselves is very clear in this table.

Table 100 - Who decides how much time a girl spends with friends				
		Girl decides	Joint decision	Adult decides
PCG	Control	50%	6%	44%
	Intervention	62%	6%	32%

The differences again appear in the last table on the decision-making issue and again are significant ($p=0.02$). Significance is conferred partly because the proportions are different but also because of the large numbers being examined by the tests. Control n=637 and Intervention n=252.

The individual results may not mean a great deal but the fact that all four are very strongly significant suggests that Intervention girls are taking more decisions than the Control girls or that they feel that they do. This does contribute to the other findings that the girls are gaining in confidence.

Life Skills Index

The Baseline Report did not make use of a Life Skills Index partly because the results from the questions in the Household Surveys were disappointing both in terms of setting baseline observations and for setting targets for future evaluation events. Some of the questions were rejected after piloting and others were used but returned virtually 100% identical responses. There is no guidance on using the Life Skills Index in either the Baseline or the Midline Report Templates – in fact, the phrase does not appear in either document.

We tried to improve the questions in the Household and Girls Survey in time for the Midline Survey and increased the number of questions on Life Skills. Again we are disappointed by the returns on a number of questions – in some cases because they produce almost unanimous responses and, in some cases, because it was not possible to detect any changes in frequencies of responses since Baseline. We are proposing another review of the Life Skills questions before the second Midline review and will attempt to base the questions on a coherent set of components of Life Skills so that we can develop an indicator built from several scales – one scale for each component.

We have seen a Life Skills Index spreadsheet which suggests a number of questions and a partial system of analysis. The analysis seems to require a disaggregation of responses by age and by Intervention and Control and the overall sample size. We have followed this model in the following tables. We are generally very positive about scalar indicators and the notes we have received in response to earlier versions of this report suggest a single score can be extracted from the Index.

We have tried to follow the model suggested by the spreadsheet but have arranged different sets of questions and their answer frequencies by putting Baseline and Midline sets together to allow easier comparisons.

Table 101 – Life Skills Index 1					
		GDM 1.	GDM 2.	GDM 3.	GDM 4.
Baseline	**Decision Making** (*I decide* + *Joint decision* %)	Whether or not you go to school	Whether or not you can go back to school or vocational training	What type of job/future career you will do	How often you spend time with your friends

		C	T	C	T	C	T	C	T
	Under 12	25%	37%	26%	31%	70%	77%	55%	57%
	12-17	23%	34%	26%	30%	68%	78%	50%	61%
	18+	27%	38%	27%	35%	82%	80%	59%	72%
	N	320	768						
Midline		GDM 1.		GDM 2.		GDM 3.		GDM 4.	
	Decision Making (*I decide* + *Joint decision* %)	Whether or not you go to school		Whether or not you can go back to school or vocational training		What type of job/future career you will do		How often you spend time with your friends	
		C	T	C	T	C	T	C	T
	Under 12	25%	21%	25%	19%	71%	68%	45%	51%
	12-17	53%	51%	42%	50%	80%	89%	54%	67%
	18+	58%	71%	54%	71%	83%	94%	88%	74%
	N	196	593						

Table 102 – Life Skills Index 2

Baseline	Making responsible choices	Going to school <u>does</u> influence what you end up doing in adult life. (Agree %)		If I do well at school, I will be able to do what I want to do in adult life. (Agree %)		The right time to get married is: *After finishing school* + *After starting work* %	
		C	T	C	T	C	T
	Under 12	87%	90%	96%	99%	77%	70%
	12-17	91%	87%	98%	98%	70%	74%
	18+	86%	89%	100%	98%	68%	66%
	N	321	769				
Midline	Making responsible choices	Going to school <u>does</u> influence what you end up doing in adult life. (Agree %)		The right time to get married is: *After finishing school* + *After starting work* %		I recognise when choices I make today can affect my life in the future (Agree %)	
		C	T	C	T	C	T
	Under 12	98%	97%	75%	62%	100%	95%
	12-17	95%	97%	79%	72%	95%	92%
	18+	94%	94%	72%	76%	92%	96%
	N	196	593				

Questions on Making responsible choices are particularly disappointing with so many of them returning undifferentiated results between 90 and 100%. After Baseline the EET informed the project about the girls who were in project schools and appeared to disagree with the statement that going to school had an influence on later life. When the project interviewed some of these girls, they were all very clear that they thought that school was important and they suggested that they had misunderstood the question. This would mean that the proportions were in fact even closer to 100% agreement with the statement.

It is possible that older girls have increased their tendency to agree with the later times to get married.

Table 103 – Life Skills Index 3						
Baseline – Self -Motivation						
	I avoid trying to learn new things when they look too difficult (Disagree %)		Life is harder for girls than for boys but there is not much you can do about it (Disagree %)		I keep trying even if others have given up (Agree %)	
	C	T	C	T	C	T
Under 12	57%	56%	27%	36%	87%	90%
12-17	60%	61%	25%	25%	91%	87%
18+	68%	70%	20%	29%	86%	89%
n	321	769				

Midline – Self-Motivation				
	I avoid trying to learn new things when they look too difficult (Disagree %)		Life is harder for girls than for boys but there is not much you can do about it (Disagree %)	
	C	T	C	T
Under 12	57%	64%	24%	30%
12-17	65%	67%	25%	29%
18+	59%	69%	15%	15%
n	208	627		

The youngest treatment girls may be saying they are more motivated now than at Baseline. Older girls may be less convinced that there are things that can be done to improve the lives of girls. This question may be too ambiguous to use in the Girls Survey. We have found it helpful in work in Tanzania on girls' education when used in a very short questionnaire. Disagreement with the statement correlates with other more progressive views on girls' education in that context.

Table 104 – Life Skills Index 4		
Baseline Problem solving	When I have a problem: *I usually ask a trusted adult for help %*	
	C	T
Under 12	87%	87%
12-17	87%	86%
18+	91%	84%
N	322	769
Midline Problem solving	When I have a problem: *I usually ask a trusted adult for help %*	
	C	T
Under 12	90%	84%
12-17	74%	75%
18+	79%	69%
N	196	593

The range of answers to the question on how to solve a problem include:
I normally work out how to solve it on my own; I ask a friend to help me; I ask a trusted adult to help me; I do not know how to solve it. The baseline results are intensely focused on relying on an adult

and show no movement with the age of the respondents. However at Midline the results show a reduced dependence on an adult as the girls get older which is probably what one would expect. The difference in highly significant.

Table 105 – Life Skills Index 5						
Midline Self-confidence	I get nervous if I have to speak in front of a group of people my age (Disagree %)		I get nervous if I have to speak in front of an adult I don't know (Disagree %)		I find it easy to tell people what I am thinking (Agree %)	
	C	T	C	T	C	T
Under 12	56%	66%	39%	30%	84%	67%
12-17	69%	67%	44%	43%	61%	70%
18+	56%	64%	42%	54%	75%	68%
n	196	593				
Midline Interpersonal skills – relationships	I can work well in a group with other people (Agree %)		I have trusted friends I can talk to when I need to (Agree %)		I have trusted adults I can talk to when I need to (Agree %)	
	C	T	C	T	C	T
Under 12	90%	89%	90%	83%	94%	90%
12-17	89%	93%	88%	83%	91%	93%
18+	100%	95%	79%	87%	100%	96%
n	196	593				

Relationships is another area where the responses seem to tell us very little as most frequencies are extremely high. There are also no signs of older girls being more confident in their relationships than younger girls which makes it unlikely that any use will come from repeating the questions.

Soft and Hard Life Skills

The Life Skills Index is almost exclusively focused on soft skills including those mentioned in the tables above: managing relationships, interpersonal skills, confidence, ambition, self-awareness, problem solving, self-control and decision-making. The project focus on hard skills includes numeracy, literacy and IT skills. Numeracy and literacy are dealt with rather extensively in Chapter 3 - Learning Outcome. The data from Learning Tests are aggregated and examined by a wide range of tests designed and mandated by the Fund Manager.

IT skills have been assessed largely in terms of how much students enjoy the lessons and what they report of their learning in their iCan journals. The reporting by the project is universally positive with all children enjoying the lessons. We have suggested developing a scalar indicator based on a range of defined skills or tasks that would be verified with a sample of students. This remains to be developed for the next reporting period. It will also be necessary to review how the new indicator is deployed as reporting has been based on those taking part in visits of the IT truck and IT training is diversifying away from this focus. These discussions have been ongoing and are not a specific feature of this evaluation – that is, there is no recommendation from the EET to the project on this monitoring work in this report as the project has already agreed on the need for change and is advancing the process. Like all scalar indicators, we expect it to be able to detect changes in a range of skills and deliver an aggregate score which can improve throughout the life of the project and the learning trajectory of the beneficiaries.

Life Skills Score

Our reporting on improvements in life skills includes a range of qualitative and quantitative observations. The Open Qual observations from individual and group interviews with girls have been particularly important in our assessments. The more quantitative findings from the Household and Girl surveys have been presented in this chapter and show some weaknesses which prevent forceful conclusions being drawn in several of the domains of soft life skills.

However, in the area of agency in decision-making the findings seem conclusive and suggest that GEC girls are more involved in decisions made about their life choices than they were at Baseline and more than Control girls. It seems appropriate to create a score based on a nominal score of zero at Baseline and improvements adding up to a score of 2.

Although we do not have ideally comparable data between Baseline and Midline it seems possible to draw from the data a conclusion that GEC girls are more confident in relating to others than Control girls. We can adopt the same approach as with decision-making and assign a nominal zero to Baseline and add a score of 1 to this component of life skills.

These scores do not prevent us from building more thorough scalar assessments for a later evaluation event. It will be possible to add other components with their own scales and aggregate the data across scales to provide a score.

Table 106 – Life Skills Index Score

	Baseline	Midline
Decision making	0	2
Confidence	0	1
	0	3

Project Checks on Intermediate Outcomes
Ensure that the IO analysis reflects the links between different levels in the logframe and informs the validity of the Theory of Change. This includes checking whether they have:

- **Measured and analysed all IO indicators presented in logframe;**

Yes – all 4 IOs were analysed

- **Disaggregated the data according to the logframe;**

Attendance – This was disaggregated by grade, not by age, as all disaggregation has been requested by grade rather than by age. We will update this in the logframe.
It was not disaggregated specifically by disability as those with complex needs have been excluded from the evaluation sample and the numbers of girls with disabilities identified in the midline is minimal. Having said this, there was analysis undertaken to find various differences and significant variations were found by CLC/type of school, which are helpful in us moving forward in how to help learning institutions collect quality attendance data.

School Management – this was disaggregated by government and private with no significant differences found, except for corporal punishment which had a highly significant difference between government (64%) and private (51%), which is highly significant statistically.
The project obviously had detailed data of progress being made in each school and can analyse this further.

Teaching quality – We had expected to disaggregate by gender, but this does not seem to be particularly relevant to the small sample we have of pay-rolled teachers. We intend to analyse by teacher, subject specialism and school when we have time after the midline submission.
Disaggregation was done by Essential criteria and teachers' self-perception of growth. We again need to dig deeper into this.

Life skills – The analysis disaggregated in various ways and found no significant differences.

- **Used both the qualitative and quantitative analysis stated in the logframe;**

Yes

- **Related the IO analysis to the analysis of Outcomes.**

Yes. Again, as a project, we need to dig deeper into what is making the biggest impact since many of the analyses has not brought out significant differences within the intervention

7 CONCLUSIONS

The Beneficiaries

The beneficiaries of the project are the girls who were contacted during GEC1. The sample of these girls that was selected at Baseline make up the Midline sample wherever we have been able to find them. Other children benefit from the work of the project, for example, where teaching quality improves all the students in the class which has received better teaching can benefit.

We analyse the data of the girls where we have data from both Baseline and Midline so that we can use pairwise assessments – effectively a *before-and-after* approach. This means that we need to track the individual girls and re-contact them at each evaluation event.

Tracking and confirming the identify of beneficiary girls are more difficult tasks than we had anticipated. More girls move more often than we had expected. There are very few fixed and certain bits of information that confirm the girls' identity. It is sometimes necessary to examine a number of different bits of information in order to assess whether the girl is the same individual who was interviewed 18 months earlier during Baseline.

The process of tracking and identifying is likely to become more difficult and more onerous as more girls move into an increasing number of schools. GEC girls have already moved into 20-30 schools where the project is not working. They are therefore not genuine beneficiaries.

The CLC girls form a special group of beneficiaries because attending a CLC is the most intense and complete support package provided by the project. It is also a well-targeted support and there is a growing population of girls who have benefited. It would be appropriate to treat this group as a special case for study. This has not been done sufficiently well up to this point.

The other GEC girls benefit from being in a school which receives a wide range of inputs from the project. These include training for school staff in management; training for teachers; access to IT equipment; access to a range of books and other reading material and support in music, games and inter-school competitions.

Other beneficiaries might be girls who are helped to attend school directly or through their families by income generating schemes or by joining a savings and loans group. The school they attend may not benefit from other project initiatives but the girls' attendance may improve.

The GEC girls therefore do not form convenient groups from which average rates of progress or scores in learning tests can be calculated easily. The relevant cross-cutting categories are based on school grades and these are not as easy to use as the strategic unit as schools or age-sets might be. The projects who have the most simple job of assessing impact are those where the school can be seen as the strategic unit and improvements in the girls' performance can be fairly attributed to the activities carried out in the school.

Non-GEC girls and boys may be beneficiaries by attending the schools where the project is working to improve management, teaching and adding to the range and quality of activities at the school. The question of including boys as beneficiaries has been discussed many times and a solution has been reached that allows the project to experiment with some initiatives that include boys as beneficiaries. This does not require much change to the project activities but it requires a significant shift in monitoring and evaluation.

Siblings of a GEC girl and other children in the same household may also be beneficiaries from changes in the attitudes and behaviour of the head of the household and the primary care giver and from improvements in the overall level of wellbeing created by involvement in income generating activities and participation in savings and loans groups.

The EET has experimented with the categorisation of the beneficiaries into subgroups although the initial recruitment of GEC girls means that the families tend to be in lower wellbeing categories in their community. Some characteristics used in creating the subgroups may be overlapping where some criteria are symptoms of others. For example - families with more disposable income may be those with more supportive approaches to girls' education but perhaps both these criteria are linked to the Head of Household having been successful in their own education.

Overall, we are comfortable that the subgroups we have created on the basis of the main source of income to the household and the highest level of education achieved by the head of the household are functional for some analyses. The other criteria which seem to link to better performance in learning tests including variables like reading at home; frequency of making savings and others may be the results of higher levels of material wellbeing. However, since the difficulties of paying fees is the most persistent and powerful of barriers, it seems appropriate to look at levels of wealth as a way of subdividing the beneficiaries.

We have seen in Devries *et al*[92], a number of questions that they used to create wellbeing subgroups based on the sleeping arrangements of the child including – *How many other children share the space you sleep in*? and, *How many adults share the space you sleep in*? we would like to test these questions in a future HHS to see if the responses corroborate our findings on subgroups.

The beneficiary households are probably poorer than the Control families. The evidence is not very strong but includes main source of income and level of education of the head of household.

	Control	Intervention
HoH main source of income is professional job	12%	6%
HoH completed secondary school or higher	16%	7%
HoH has none or little Primary schooling	39%	50%

There are other observations on the number of school-age mothers and brides in Intervention, although very small, is many times the number in Control. And the data on transition of in-school girls show more variability among the Intervention girls than Control girls perhaps implying the latter have a more robust experience in education. See Annex 3 for a more detailed discussion of this issue. Overall, we judge that the Control may be too hard a test for the Intervention girls and the differences shown in the Learning Tests may be an underestimate of the effects of the project activities.

A new approach to sampling and following girls' progress may be necessary to avoid problems of tracking and identifying GEC girls; the difficulties of comparability between Intervention and Control samples and being able to attribute observed changes to specific project activities.

The Context

The barriers to girls' education remain largely unchanged and the financial cost of education is the main issue cited. The situation is not simple and the project response, consisting of a wide range of initiatives targeting different components of the barriers, may be an appropriate response although very difficult to monitor and evaluate.

[92] *Op.cit*. see Table 4, p7.

One question in the Household survey has thrown some light on reductions to the principal barrier of having to pay fees. Some PCGs mention that it has become easier to find the necessary money and others have found that school staff have become more accommodating in the timing of delivery of the fees. The second most common change that has led to easier access has been the attitudes of the girls who have become more keen to study and seem to like school more. We believe better teaching methods and less use of corporal punishment have been instrumental in bringing about these changes in attitude. Parents in the survey qual work and men in the Open Qual say that the project initiative called Positive Parenting has been largely responsible for changes of attitudes among PCGs.

The project has modified its portfolio of activities in the period between Baseline and Midline introducing more income generating activities and exploring the possibilities of work for boys. Work in Karamoja has increased with a Mentor being stationed there. Interviews with

The project Theory of Change is hard to follow and a more comprehensible and perhaps less complete version is in preparation. Clearer statements of the project objectives would be welcome.

Levels of violence against children may be higher than reported in previous evaluation and monitoring work, especially in urban areas. The evaluation responses show higher levels in urban areas and in all the questions on safety in the Household Survey a minority (5-10%) show they are worried about safety. If this is a serious underestimate of the scale of the problems it is possible that the EET survey methods, including interviews when many people are around on a school premises, may not be the best way to elicit responses in this difficult area.

Learning Tests

Intervention girls have made more progress in Learning Tests between Baseline and Midline than Control girls. The extent of their greater progress is not as much as the target set by the Fund Manager. The progress made is not uniform in all grades but some grades, especially those where the project has focused its efforts, show significant differences between Intervention and Control.

The Intervention girls make more progress than Control girls in every grade and so even though some of the differences are not significant, the overall story is inescapable.

The EET used exactly the same model of Learning Tests at Midline as at Baseline. This may not have been a wise approach since we knew that many Early Grade subtasks showed strong ceiling effects, but our feeling was that the best way of demonstrating changes between Baseline and Midline was to use exactly the same tests.

The mean marks for different grades show gradual increases with the age of the girls. The changes plateau after the girls reach 15 years old and increasing the weight of the Senior Grade tests did not completely eradicate the plateau. The standard deviations of the means are usually large and often larger than the difference in means between one grade and the succeeding year.

After the Baseline learning tests, a girl will have spent three terms in her previous grade and one term in the grade she is in at the time she sits the tests at Midline. Girls are identified by the grade there were in at Baseline.

A great deal of time and effort are put into designing, piloting and refining the learning tests; training enumerators to administer the tests and carrying out the tests with the girls and the

management, cleaning and analysing the data from the tests. This absorbs most of the resources allocated to the evaluation team. The learning for the project is very slight. The girls' teachers do not see the tests and the girls do not find out their marks. The level of ethical jeopardy in managing tests in this manner is high.

The Learning Tests show increases in mean marks by grade up to S3 where the changes level off. It may be possible to eliminate the plateau by changing the weighting to increase the influence of the SeG tests on the aggregate marks. This can be assessed as part of an overall revision of the Learning Tests for future evaluation events.

Marks in literacy correlate with marks in numeracy when the data from girls in school from P4 to S5 with both Intervention and Control girls taken together (r=0.79, n=624). This implies that girls who are good at literacy tend also to be good at numeracy. Of course, this is not always the case but it seems sufficiently frequent to allow the generalisation to be made and it suggests that the aggregate learning test marks are assessing something of the competencies of the girls in literacy and numeracy. The distribution of aggregate marks in the learning tests conform to a Normal Distribution which would suggest that the tests are appropriate despite the presence of ceiling and floor effects in the individual subtasks. The normal distribution could itself be an artefact and this can be tested when the Midline 2 tests are being designed.

The target set by the Fund Manager appears to be arbitrary and Miske and Joglekar report that the literacy target set in GEC1 was not explained in any of the FM's guidance materials and add, "*Clearly it had never been tested as a target in a PbR intervention*". They go on to say that the pre-set target, "*must be reviewed carefully*[93]". This is partly because it is counterintuitive to expect literacy to improve according to a linear trajectory, something the authors find no support for in any of the studies they identified.

The differences between Control and Intervention averages in any one grade tend to be smaller than the differences between different grades. It may be that improved teaching in improved environments with more motivated girls cannot make as much difference as an extra year of schooling.

It is clear from the other areas of the survey work and the qualitative interviews that a lot is working better because of the project initiatives. The girls interviewed in the Open Qual work say they are working better and enjoying school more. The Learning Support Teachers and Headteachers in the survey qual work say the same and that the better teaching methods and the reduction in corporal punishments are creating happier school environments. The better teaching methods form a long list and most of them are praised for making the lessons more inclusive and the teaching more appropriate for the range of learners in the lesson. The girls have more opportunities to speak and interact according to their own views (Open Qual) and those of the Learning Support Teachers (survey qual). This may feedback on their declared levels of confidence (Open Qual and survey qual) which would lead to better performance in Learning Tests.

But the improved teaching and confidence and other changes do not lead to higher mean aggregate marks in the way required by the Fund Manager. It may be that the changes in performance in Learning Tests only appear after a longer period of improved inputs. The impact chain may be longer than the time elapsed between Baseline and Midline. And it may be that the differences between Intervention and Control are not sufficiently marked because the Control girls are from slightly better-off households and have a longer more

[93] Shirley J. Miske, Alison B. Joglekar, "*Using or Misusing the Early Grade Reading Assessment? Examining A Measure of Payment by Results in the Girls' Education Challenge*" In Annual Review of Comparative and International Education 2017. Published online: 22 Jan 2018; 187-202 (p199).

consistent education history from the Intervention girls. There is some evidence for this (see above Conclusions – The Beneficiaries).

But it may simply be that the tests are not sensitive enough, or to put it another way, girls' performances in the learning tests are too variable. The variations about the means are large so it is harder to find significant differences. The variability may be due to the unfamiliarity with the tests and the evaluation event which must appear like a disruption to a normal school day. The fact that some girls did not complete the tests suggests that they were not clear on what was expected of them or perhaps chose to avoid doing the tests. It is also true that when the subtasks are examined girls do less well in later tests despite presumably being more capable. This is common in the tests done in the CLCs and must be due to the level of attention given to the task by the girl rather than her actual level of competence. It seems likely that a shorter survey interview and shorter and fewer learning tests would lead to better levels of attention and better completion of test sub-tasks.

Transition

The EET had access to six, and sometimes eight, estimates, of the In-School Girls' grades from the Baseline and Midline survey work which enabled a much more accurate statement of the girls' grades and the changes in their grades since Baseline. The data show that almost all In-School girls progressed by one grade – an almost perfect transition. A tiny number of school-age girls dropped out from the Intervention group but this contrasts with no girls from the same grades dropping out from Control. Small numbers repeated a year which was not considered a failed transition at Baseline but is at this evaluation event.

Small differences exist between Intervention and Control when marriage and childbirth are examined. There are not many Intervention girls who are married and or have a baby but there are far more than in Control. This seems likely to be a result of the different recruitment methods to Intervention and Control samples.

The transition data in this report show the situation at the time of the Midline survey. We know that project staff are monitoring the girls who drop out and two of those of school-age who had dropped out at the time of the Midline survey are now back in school.

66% of the Intervention girls who were out of school at Baseline have since made successful transitions into school, work or training. The percentage for Control girls is lower (33%) but the overall numbers are much smaller.

The project focus on P7 girls seems to be linked with a very high level of successful transition into S1 which had previously been a key moment for failed transitions.

There are three transitions which are more problematic among Intervention girls. A high number seem to repeat P6 and we will explore why this might be. It's possible that teachers fear that the girls may not be ready for the leavers examination at the end of P7.

There are also low rates of successful transitions for girls leaving S4 and S6 although numbers are small. Both these grades are ages when girls can stop full-time school education and make a successful transition into work or training or other forms of education. However, a number seem to be recorded only as Out of School which constitutes a failed transition. This situation may have changed since the Midline survey and this may be beyond the remit and responsibilities of the project. If Intervention girls have got as far as completing S4 this may already be a success for the project and the EET cannot, in any case, compare the rates of transition with those of the Control girls since there are none of the appropriate age in the sample.

It seems hard to set targets for Transitions as the project seems to be doing as well as could be expected. A target could be achieved or missed because of the behaviour of a tiny number of girls. We have suggested that targets for those grades and situations that currently show successful transition rates above 90% should simply be set at that level. Improvements in transition rates are suggested for girls who might move from P6 to P7 and the question is posed for improved outcomes for S4 and S6 girls.

Attendance
Since Baseline, attendance observations concern only GEC girls – the project does not take responsibility for the entire population in school or training centre but only for those girls it has or has had contact with directly. Attendance at CLCs is a project responsibility. The Household Survey does not deliver very useful data on attendance.

Attendance is assessed by spot checks and where registers are well managed the overall attendance levels on any day are between 80 and 90%. The quality of data is poor and evidently there is no incentive for mainstream teachers to keep registers up-to-date. Some well-maintained registers are difficult to interpret correctly.

The project has developed a new app which can run on a mobile phone and deliver day to day reports of the presence or absence of individual GEC girls. There are management issues with the app and the returns are of variable quality. There are also technical issues with the use of the data but the approach shows great promise.

The qualitative interviews reveal that Learning Support Teachers do more work on contacting the families of GEC girls who do not arrive at school. This may be partly due to using the attendance app which may highlight absences that might otherwise not be noticed. Absentees are followed up by the Mentors or the LSTs themselves. PCGs say that they feel contact with the school has improved even though in general they visit the school less often than they claimed at Baseline.

The biggest improvement in attendance has been because paying fees has become less of a barrier for some girls. This is partly due to increased income on the part of the PCG and partly because the schools have shown themselves better able to be flexible and, for example, to accept payments in instalments. The financial barrier remains the most important overall.

Attendance has also become easier because of improvements in the attitudes and ambitions of GEC girls.

The attendance Intermediate Outcome highlights issues with the project approach of focusing on the individual girls at the expense of following what is happening within an institution or a cohort of girls. The approach makes sense in terms of the requirement from the Fund Manager to work only with the GEC girls and where the project can realistically take responsibility for work that benefits the girls. It is not clear if this is sustainable since the girls are very mobile and the work of tracking is very difficult and requires a lot of resources. It is probably impossible and certainly impractical to assess the attendance of GEC girls where there are only a few in a school. The continuing dispersal of GEC girls will make this an increasingly difficult problem.

Better questions are necessary in the Household Survey that will produce better information on rates of attendance and reasons for increasing or decreasing attendance.

Teaching Quality
It is not always easy to separate teaching quality from safeguarding issues. Teaching in a safer space is better quality teaching. Learning Support Teachers talked to us positively

about a huge range of teaching methods that they believe are raising the quality of teaching. Headteachers also spoke positively about methods that engage and involve the pupils so that they are becoming more confident. Girls in a different set of interviews talked about similar classroom teaching methods and also about being engaged in activities during school assembly and in clubs which have helped them to be more confident and outgoing.

Some rather nuanced reasoning is required to make sense of two observations in which PCGs' responses in the Household Survey appear to suggest negative trends in teaching quality. PCGs appear to have become more critical of teaching methods and agree more at Midline than they did at Baseline with statements that teachers have sexist ideas about the competencies of girls and boys and focus more on the better pupils. However, all PCGs say that the overall quality of teaching has improved since Baseline. The first observations may be of raised awareness of the issues of gender bias and issues of inclusive teaching for all pupils.

The PCGs in Intervention areas believe that teachers use more imaginative teaching methods that those in Control areas – the difference is significant. Learning Test results are better where fewer girls agree that the main method of teaching is for pupils to copy from the board.

The project methods of engaging with teachers and promoting experimentation and new approaches seem effective and positive. Lesson observations are providing monitoring data based on some essential and desirable criteria that should be present in good lessons. Teachers are making progress in this monitoring. The difficulties of teaching with few resources in very over-crowded classes are easy to underestimate. Learning Support Teachers mention difficulties in teaching with some teachers who do not have progressive ideas on how teaching should be done.

Girls think that teachers being absent from the classroom is a more serious problem that the carers (very significant difference). More work is required on exploring issues of absenteeism among teachers. Reported levels in the Household Survey seem lower than the levels published in other work on this issue.

School Management

The project QIS training and assessment programme is delivering good monitoring data on the progress made by different schools in different domains. The QIS data when examined by each of the indicators in each of the aims show very important improvements in all schools. There are greater improvements in areas where the project has placed more emphasis.

Child Protection has been an important area and shows the highest rate of improvement. . In a different exercise Mentors suggested that work on CP in schools had not been very successful probably because they have revised their optimistic assessments made at Baseline. The EET judges that the two findings corroborate each other and that child protection is improving in schools where the project is working. This is another case where lower rating is taken to mean better performance because the people doing the rating have clearer ideas on what constitutes progress.

The Household Survey results suggest that PCGs think school management is improving with increased numbers (15% at Baseline up to 25% at Midline) saying that the school is managed "*extremely well*"; increased numbers saying that the Headteacher is doing an "*excellent*" job and approving the presence of a Child Protection Policy. Nonetheless, there are still 5% of respondents who feel that girls are not safe at school. On the question of reporting all cases of abuse the Intervention area carers are significantly more confident than those in Control. But both Intervention and Control parents and carers are less confident

that all abuse is reported than they were at Baseline. Again, this is likely to be due to raised awareness. School Management overlaps with Teaching Quality and safeguarding in this area.

Life Skills
The Open Qual interviews (see also under Teaching Quality) contained a great deal of positive reporting from the girls on their raised confidence and ambition. They talked a lot about getting skills, having better relationships with their parents and with their friends and not being "jumpy" like some girls are around boys. The parents in their individual and group discussions talked in the same way mentioning working in new initiatives (16 new activities were mentioned in two group discussions) and listening more to their girls.

The Household Survey results corroborate some of these ideas. GEC girls are significantly more confident in speaking in front of children their own age and in front of adults they don't know. All girls are significantly more confident in facing difficult situations and less likely to give up trying but this may be a result of getting older. The Survey asks four questions about decision making and in all cases the intervention girls are significantly more likely to make the decisions themselves than the girls in Control.

The story from the Household Survey and the Open Qual is the same – GEC girls are more confident and have greater ambition than Control girls.

The other questions in the survey on life skills are disappointing and provide little learning as nearly all respondents choose the same response. Although one can conclude that positive attitudes have not altered between Baseline and Midline, there is in fact little to gain from this observation. Some new questions are required in the Household Survey that will identify areas where there is likely to be some observable changes in attitudes, confidence and ambition.

Sustainability
The Sustainability Scorecard is working well as a method of collecting the views of the Mentors on changes that are happening in local communities and schools. Overall the progress is good and the reporting makes sense generally corroborating the findings in for example monitoring in School Management and Teaching Quality. In both areas the observations of greater progress coincide with the areas where the project has invested more effort.

There are nice paradoxes around Child Protection where the Mentors are less convinced about progress and even suggest negative change at community level. The discussions reveal that the Mentors are revising their opinions of the situation at Baseline based on greater awareness now.

Engagement between schools and parents is seen to be better than at Baseline which contradicts the findings from the Household Survey in which parents say they are visiting the school less often and hearing information from the school less frequently. As mentioned above, the difference is probably due to Learning Support Teachers and Mentors being in touch directly with PCGs. Community leaders tend to lag behind the PCGs in their thinking about girls' education.

The Open Qual again provides ideas on sustainability from the numbers of positive comments from carers and girls on new activities, greater income, more confidence and more ambition. These observations corroborate the finding from the HHS question on why has access to school become easier where the second most important change after the reduction in the financial barrier is the improvement in girls' attitudes to school.

A father who took part in the Open Qual interviews explaining that a lot of people still feel negatively towards girls' education said that he hadn't heard of a girl going to university until 10-15 years ago and much more recently than that he heard for the first time that a girl had got top marks in her year in school. These remarks help to explain that it is a long way from changing the attitudes of the carers of the GEC girls to changing social norms around key cultural issues.

GESI

Social inclusion starts with the selection of GEC girls based on their difficulties with schooling which equates to targeting the poorest families in the community. The focus on poorer and more disadvantaged girls informs the activities and because the initial selection has determined the beneficiary population the targeting is not called into question and monitoring does not disaggregate results by levels of wellbeing. The Midline Evaluation has identified more observations including the numbers of girls getting married and having babies which are consistent with the GEC beneficiaries being poorer and less secure in education than other families in their communities and also in comparison with the Control population.

The gender approach of the project comes from a good analysis of the current situation which is well described and discussed in the gender framework. The work is designed to address the key elements of the how girls are impacted because of their gender and the aims are to improve their situation in a sustainable way effectively reducing the gender gaps in access to education, access to information technology and to the other benefits of being in education. The gender framework could be updated and contain more areas of assessment. Reference could be made to more recent reports on the State of the World's Children[94] and the special report on children with disabilities[95]. A new version could also include more up-to-date information on violence against children and the key points shared in special staff meetings so that all staff have access to similar information and learning.

The work with the Karamojong girls supported by one of the partners is a special case of social inclusion because of the levels of disadvantage and discrimination faced by people in this indigenous group.

Working with girls with disabilities is leading to some significant successes in social integration with small numbers of girls being able to join and remain in mainstream school and to attend special schools. Others are becoming included through economic activities and through changed attitudes among members of their family and neighbours. The numbers are small but the impacts on the individuals is important.

The household survey results suggest that relatively minor difficulties ("some difficulty[96]") may affect performance in learning tests. It may be useful to investigate some of the cases that emerge from the survey to see if the girls' performance in the learning tests reflects having more difficulties at school than the level of disability might imply.

Not working for boys has been a high-level strategic decision that appears to have been changed. Introducing more work with and for boys will require additional funding and some skill so that the work reinforces or supplements the work being done for girls. It could take the form of support through CLCs based initially on boys from the families of the GEC girls. This might mean supporting the families more thoroughly and doing something with boys that has been proven to be effective with girls and not extend the project work too much.

---

[94] https://www.unicef.org/sowc/

[95] https://www.unicef.org/sowc2013/

[96] "Some difficulty" is the second of four categories in the Washington Group scoring of disability. The FM definition of having a disability requires the score of at least "*a lot of difficulty"* which is the third of the four categories.

The Evaluation tools

This report comments in several places on the tools used by EET and those comments are clarified and brought together here.

Household Survey

The Household Survey is too long and is not focused tightly enough on the key issues confronting the work of the project or the key questions asked by the Fund Manager. Routing has proved an issue in correctly identifying girls who are genuinely Out of School that is the girls who are of school age, not in school and not in work or training. It has also not allowed easy analyses of the CLC graduates. The Life Skills area seems particularly weak despite producing a few crucial statements of project impact. Too many questions are transparent allowing respondents to choose the "right" answer and leave us with 99% agreement with a positive-sounding statement. Some questions have not been exploited at all and should be removed if there is no possibility of them being used in future surveys.

Learning Tests

The Learning Tests are too long and contain subtasks that have high ceiling effects and on their own do not tell us enough about progress made by students. The subtasks seem to jump from low level tasks to much higher-level tasks which discourage the girls from continuing. The administration of the tests needs to be stricter – too many girls failed to take part in all the necessary tests.

The Early Grade subtasks should be examined for ceiling effects and those where the effects are serious should be removed. The possibility of using of Oral Reading Fluency as the only EGRA task should be tested by seeing if early results show a correlation between the overall aggregate marks and the ORF score.

The Learning Tests for Midline 2 will need to be checked for errors and the questions will need to be arranged in order of increasing difficulty. Some pretesting will be necessary and useful.

New ways of calculating overall aggregate marks in literacy and numeracy must be developed which will be easier to calculate. The results will need to allow continuity from the Baseline and Midline results and a continuation of testing through to the end of the programme.

Qualitative

The Open Qual interviews have provided very important observations and the exercise should be repeated before Midline 2. The qualitative survey interviews have been invaluable and another round of interviews should be carried out in 2020 in order to test some of the findings from the Midline surveys. Interviews with CRANE staff should be repeated regularly.

8 RECOMMENDATIONS

A New Strategy
The project has been obliged to work following the GEC girls and this has worked well but required huge resources in tracking and identifying the girls and their families at the Baseline and Midline evaluation points. The work will get harder as the girls continue to disperse to more different locations and more different schools or other institutions and activities.

Tracking girls individually makes it difficult to assess the quality the institutions they attend and the impact of some of the project initiatives. At the same time, focusing on the GEC girls prevents assessments of the impact of project work on other children in the same schools and in the same families as the girls.

The requirement to deliver learning tests results by grades creates a mismatch with the focus on individual girls and their learning histories. It is time for a review of the entire project strategy with respect to following changes in levels of literacy and numeracy. This is referred to in the text of this report as a Post-Midline Review.

The project needs a simpler operational model that is not based on working for all the GEC girls and only the GEC girls. This not only makes evaluation difficult and different from most monitoring but also neglects reporting where project work was having an effect on other children.

If a more coherent model can be developed, it may make some of the following recommendations less important. It would be interesting for example to maintain a tracking approach for CLC girls and continue to monitor their progress and transitions and, where the girls are in school, their progress in literacy and numeracy. At the same time a cohort approach could be used in a small number of schools where entire classes could sit the Learning Tests and sufficient numbers of cases would be examined to allow changes in aggregate marks to be assessed. Changes might be related to specific work done in the school. If the Learning Tests are carried out in school, the results could be shared with the teachers and the girls which might lessen the ethical issues in testing.

Work should be continued to identify and start initiatives for boys. This work is likely to start at a small scale and should be based on learning from the history of the project. It might be a mistake to launch totally new and untried areas of work.

The simpler theory of change that is in draft should be revisited after the new strategy is planned. It should aim for simplicity and clarity at the expense of describing every project activity and every link between activities and results.

The Household Survey
The EET did follow its own recommendation at Baseline to improve the routing of the Household Survey and to improve some of the questions. Some of the questions that appeared at Baseline to be weak or not to deliver useful learning were allowed a second chance to prove useful at Midline. With the additional learning from Midline we can now remove large numbers of questions that do not help our understanding of the situation or of changes in the situation. The survey is too heavy and can be reduced in scale.

The Household Survey should be modified so that it is simpler and easier to administer and analyse. The routing should be improved to allow better learning about the situation of CLC girls and disentangle them from other girls who might be seen to be out of school. The section of Life Skills needs editing to remove redundant questions and test some that will

provide learning about levels of confidence and ambition. Better questions should be trialled to assess levels of material well-being.

The Learning Tests
The Learning Tests need to be examined in detail for the learning they are providing and how the different subtasks correlate with each other so that there can be a thorough overhaul of the subtasks and a much lighter set of tests can be used in future evaluation events. The Midline tests were reworked considerably from the draft material produced at Baseline. The third set of tests will need editing and checking before they are used. This will be easier if some subtasks can be eliminated from the overall testing.

The cost of Learning Tests
The EET, the project staff and the Fund Manager should consider ways to reduce the demands of the Learning Tests while allowing monitoring of girls' performance in literacy and numeracy to continue. This might include reducing the size of the sample (fewer girls); reducing the number of components in the Learning Tests (fewer subtasks) and more time between evaluations (fewer surveys).

The EET could provide better evaluation work which strengthened and cross-checked the project monitoring if it were released from some of the pressure currently applied to delivering on the Learning Tests. The time between this Midline review and the next evaluation event is the time to review the focus of the project and the design of evaluations.

Gender expertise
It would be good to update the Gender Framework with some more recent references and a simple overall set of statements that outline the situation in education and the project responses.

The new Gender Framework should be presented to all staff to help support learning and a shared understanding of the context that the project is working in and how its work addresses the ways people are affected because of their gender.

The presentation of the updated Gender Framework could be broadened into a more general Gender Training event. This need not be long or costly but would help to build a deeper understanding and a shared analysis and vocabulary for discussing gender issues.

Assessing Sustainability
The Sustainability Scorecard is providing interesting assessments of important changes in different domains. However, it seems to be poorly integrated with the other monitoring that is being done routinely. The value of the Sustainability Scorecard activities is drawn out of the conversations and sharing of information and ideas that occur as the participants debate the scores that they want to attribute to the different indicators. It makes sense, for example, to take learning from the monitoring of teaching quality into the discussions of Indicator 10; learning from the QIS work into conversation about Indicator 12 and call staff involved in working with the Ministry of Education to discuss Indicator 17. It is not normal to have different indicators assessing the same changes but this can be built on to strengthen the links between monitoring and evaluation.

Transport and sustainability
It would be good to put some other models of transport to the test given the difficulty of maintaining the library, IT and transport buses. Some models of shared taxi use and collective bargaining for lower costs could be examined. Different methods of moving books between schools and promoting reading or promoting the use of IT equipment and giving computing lessons that don't require a large vehicle. Perhaps collectives made up of five or six schools that could somehow share the costs of moving a book bag from school to school.

The initial arrangements could be subsidised by the project until the real costs are clearly costed and the potential examined. The probability of long-term success might not be good but ideas like these or others from people who know the situation better than the EET would be likely to continue longer than the vehicles after GEC-T.

Attendance
The monitoring of attendance is potentially important but seems difficult to manage and schools do not collect or keep adequately accurate or up-to-date data. The phone-based app has great potential and should be deployed in a small number of locations where it can be used to deliver good quality information. There are some management issues on compliance and covering costs that may need to be sorted out.

There is a protocol issue that will clarify the meaning of a blank record and perhaps a mechanism that only allows data to be entered on the day to which it applies. The monitoring officer needs to be trained in management of data in excel so that the work of collating observations and preparing attendance reports can be carried out regularly with relatively little work. It might be possible to lighten the load by reporting on a sample of GEC girls whose results could be extrapolated to create an overall estimate of levels of attendance. It might also therefore be possible to link levels of attendance with performance which so far has not been possible.

Targeted follow up
There are more data and learning from the evaluation that could be used for targeted follow up of girls with disabilities and who appear to have dropped out from school or who could be helped with some small but targeted support. This could be reviewed by the EET and CRANE staff to identify a small number of cases to be taken forward.

Project contribution: Response to conclusions and recommendations

See Annex 17