

---

# Somali Girls' Education Promotion Project – Transition

## Midline Report

15 July 2019

Brenton Peterson, Jonathan Forney and Samuel Ha

Version 4

By



For



<b>LIST OF TABLES .....</b>	<b>4</b>
<b>LIST OF FIGURES .....</b>	<b>7</b>
<b>EXECUTIVE SUMMARY .....</b>	<b>8</b>
<b>2. BACKGROUND TO PROJECT.....</b>	<b>14</b>
2.1. PROJECT OVERVIEW.....	14
2.2. PROJECT CONTEXT.....	14
2.3. PROJECT THEORY OF CHANGE AND ASSUMPTIONS .....	18
<b>3. EVALUATION METHODOLOGY .....</b>	<b>36</b>
3.1. KEY EVALUATION QUESTIONS & ROLE OF THE MIDLINE.....	36
3.2. SAMPLING METHODOLOGY.....	42
3.3. LIMITATIONS AND CHALLENGES .....	50
3.4. RE-CONTACT RATES AND PANEL ATTRITION.....	53
<b>4. CONTEXT, EDUCATIONAL MARGINALISATION AND INTERSECTION BETWEEN BARRIERS AND CHARACTERISTICS</b>	<b>60</b>
4.1. EDUCATIONAL MARGINALISATION .....	60
4.2. INTERSECTION BETWEEN BARRIERS AND KEY CHARACTERISTICS .....	66
4.3. APPROPRIATENESS OF PROJECT ACTIVITIES TO THE CHARACTERISTICS AND BARRIERS IDENTIFIED .....	70
<b>5. KEY OUTCOME FINDINGS.....</b>	<b>71</b>
5.1. LEARNING OUTCOME.....	71
5.2. SUBGROUP ANALYSIS OF LEARNING OUTCOMES .....	103
5.3. TESTING THE TOC – INTERMEDIATE OUTCOMES AND LEARNING OUTCOMES .....	113
5.4. TRANSITION OUTCOME.....	116
5.5. SUBGROUP ANALYSIS OF TRANSITION RATES.....	131
5.6. TESTING THE TOC – INTERMEDIATE OUTCOMES AND TRANSITION.....	140
5.7. SUSTAINABILITY OUTCOME .....	143
<b>6. KEY INTERMEDIATE OUTCOME FINDINGS .....</b>	<b>162</b>
6.1. ATTENDANCE .....	162
6.2. SCHOOL GOVERNANCE AND MANAGEMENT .....	184
<b>6.3. QUALITY OF TEACHING .....</b>	<b>194</b>
6.4. LIFE SKILLS.....	214
6.5. COMMUNITY-BASED ATTITUDES AND BEHAVIOUR CHANGE .....	230
6.6. SCHOOL-RELATED GENDER-BASED VIOLENCE.....	234
<b>7. ALP GIRLS BASELINE.....</b>	<b>237</b>
7.1. PROFILE OF ALP GIRLS.....	237
7.2. ALP GIRLS’ LEARNING OUTCOMES.....	246
7.3. ALP – SUBGROUP ANALYSIS OF THE LEARNING OUTCOMES.....	252
<b>8. CONCLUSION &amp; RECOMMENDATIONS .....</b>	<b>255</b>
8.1. CONCLUSIONS.....	255
8.2. RECOMMENDATIONS.....	261
<b>ANNEX 1: MIDLINE EVALUATION SUBMISSION PROCESS.....</b>	<b>265</b>
<b>ANNEX 2: INTERVENTION ROLL-OUT DATES .....</b>	<b>266</b>

<b>ANNEX 3: MIDLINE EVALUATION APPROACH AND METHODOLOGY .....</b>	<b>267</b>
<b>ANNEX 4: CHARACTERISTICS AND BARRIERS .....</b>	<b>276</b>
<b>ANNEX 5: LOGFRAME .....</b>	<b>277</b>
<b>ANNEX 6: OUTCOMES SPREADSHEET.....</b>	<b>278</b>
<b>ANNEX 7: PROJECT DESIGN AND INTERVENTION .....</b>	<b>279</b>
<b>ANNEX 8: KEY FINDINGS ON OUTPUT INDICATORS.....</b>	<b>283</b>
<b>ANNEX 9: BENEFICIARIES TABLES.....</b>	<b>293</b>
<b>ANNEX 10: MEL FRAMEWORK .....</b>	<b>297</b>
<b>ANNEX 11: EXTERNAL EVALUATOR’S INCEPTION REPORT (WHERE APPLICABLE) .....</b>	<b>297</b>
<b>ANNEX 12: DATA COLLECTION TOOLS USED FOR MIDLINE.....</b>	<b>297</b>
<b>ANNEX 13: DATASETS, CODEBOOKS AND PROGRAMS .....</b>	<b>297</b>
CODEBOOK.....	298
KEY REGRESSION RESULTS .....	301
<b>ANNEX 14: LEARNING TEST PILOT AND CALIBRATION.....</b>	<b>318</b>
<b>ANNEX 15: SAMPLING FRAMEWORK .....</b>	<b>322</b>

## List of Tables

Table 1: Project design and intervention .....	25
Table 2: Evaluation sample breakdown of girls and boys (by region) .....	33
Table 3: Evaluation sample breakdown of girls and boys (by grade) .....	34
Table 4: Evaluation sample breakdown of girls and boys (by age) .....	35
Table 6: Targeted and achieved samples, by tool, in midline .....	47
Table 7 - Learning and transition samples .....	47
Table 8: Baseline and midline samples, by tool .....	48
Table 9: Targeted sample sizes for qualitative data collection, by tool or population group .....	49
Table 10: Re-contact rates by age and grade .....	55
Table 10: Comparison of baseline cohort girls and their replacements .....	57
Table 11: Girls' characteristics .....	60
Table 12: Evaluation sample breakdown of girls (by aggregated disabilities) .....	62
Table 13: Evaluation sample breakdown of mental health impairments of only comparable girls .....	62
Table 14: Evaluation sample breakdown of girls (by disaggregated disabilities) .....	63
Table 15: Potential barriers to learning and transition .....	64
Table 16: Migration by Girls and Boys (age 11-21) in Cohort Girl's Household .....	66
Table 17: Intersection of barriers to education by subgroup characteristics for intervention areas .....	68
Table 19: Intersection of barriers to education by subgroup characteristics for comparison areas .....	69
Table 19: Literacy (EGRA/SeGRA) .....	76
Table 20: Numeracy (EGMA/SeGMA) .....	79
Table 21: Foundational numeracy skills gaps .....	83
Table 22: Foundational numeracy skills gaps with mean scores by grade .....	84
Table 23: Aggregate DID analysis of mean numeracy scores by subtask .....	84
Table 24: Foundational literacy skills gaps .....	88
Table 25: Foundational literacy skills gaps with mean scores by grade .....	88
Table 26: Aggregate DID analysis of mean literacy scores by subtask .....	89
Table 27: Literacy (EGRA/SeGRA) .....	93
Table 28: Numeracy (EGMA/SeGMA) .....	94
Table 29: Boys' foundational numeracy skills gaps .....	98
Table 30: Aggregate DID analysis of boys' mean numeracy scores by subtask .....	99
Table 31: Boys' foundational literacy skills gaps .....	100
Table 32: Aggregate DID analysis of boys' mean numeracy scores by subtask .....	100
Table 33: Learning scores of boys' subgroups .....	101
Table 34: Grade Level Standards for Numeracy .....	101
Table 35: Numeracy grade level achieved by intervention (and comparison girls in parentheses), by Grade .....	102
Table 36: Boys' numeracy grade level achieved by intervention (and comparison boys in parentheses), by Grade .....	103
Table 37: Transition performance against targets .....	116
Table 39: Transition Pathways Matrix .....	118
Table 39: Transition rates and pathways, intervention group .....	121
Table 40: Transition rates and pathways, comparison group .....	122
Table 41: Overall transition rates and difference-in-differences .....	127
Table 42: Transition rates and difference-in-differences estimates by age group .....	128
Table 43: Transition targets for future evaluations .....	131
Table 44: Transition rates at midline, by subgroup .....	132
Table 45: Transition rates at midline, disaggregated by barriers to transition .....	137
Table 47: Sustainability Scorecard .....	143
Table 47: School-level indicators of CEC engagement .....	160

Table 48: Community-level indicators of CEC engagement.....	161
Table 49: Attendance Yesterday, Today, and of Headcount for Girls and Boys in Midline and Baseline – Headcount Survey .....	163
Table 50: Boys’ Average Attendance Rate – Headcount Survey.....	164
Table 51: Girls’ Average Attendance Rate – Headcount Survey .....	164
Table 52: Percentage Difference in the Average Attendance Rate at Midline vs. Baseline .....	165
Table 53: Girls’ Attendance in Baseline and Midline – School Survey.....	166
Table 54: Limited Girls’ Attendance in Baseline and Midline – School Survey.....	167
Table 55: Limited Girls’ Attendance in Baseline and Midline – Household Survey.....	170
Table 56: Measures of Girls’ Attendance in Baseline and Midline .....	172
Table 57: Attendance of key subgroups.....	179
Table 58: Attendance of key barriers.....	182
Table 59: Project impact (difference-in-differences) on school management and governance indicators ...	185
Table 60: Project impact (difference-in-differences) on household measures of school management .....	191
Table 61: Provision of Financial Support to students in Midline- Household Survey.....	192
Table 62: Classroom participation and learner-centred methods, by intervention status and round.....	196
Table 63: Use of corporal punishment, by intervention status and round.....	203
Table 68: Teachers’ classroom demeanour, by intervention status and round.....	205
Table 65: Teachers’ use of encouraging feedback for boys and girls, by intervention status and round .....	209
Table 66: Teacher preparedness and effort level.....	211
Table 67: Effect of teacher training on classroom demeanour, participation and the use of formative assessments.....	213
Table 68: Project impact (difference-in-differences) on self-confidence .....	215
Table 69. In-school girls’ average YLI score .....	216
Table 70: Life-Skills Survey Questions.....	218
Table 71: Reported as percentage stating “I decide” or “I and my family decide jointly” .....	219
Table 72: Reported as percentage stating “strongly agree” & “agree” .....	220
Table 73: Reported as percentage stating “strongly agree” & “agree” .....	221
Table 74: Reported as percentage stating “strongly agree” and “agree” (for girls with disability).....	222
Table 75: Project impact (difference-in-differences) on self-confidence .....	224
Table 76: Life Skills Index Difference in Difference Results.....	226
Table 77: Average Life Skills Index Score.....	227
Table 78: Caregivers’ perceptions on worth of girls’ education.....	232
Table 79: Household characteristics of ALP, in-school and OOS girls .....	240
Table 80: Economic status of ALP, in-school, and OOS girls’ households.....	241
Table 81: Views of girls’ education and girls’ decision-making power.....	242
Table 82: Life skills and self-confidence of ALP, in-school, and OOS girls.....	245
Table 83: ALP Literacy (EGRA/SeGRA) .....	246
Table 84: ALP Numeracy (EGMA/SeGMA).....	247
Table 85: ALP Foundational numeracy skills gaps .....	248
Table 86: ALP Foundational literacy skills gaps.....	250
Table 87: Grade Level Standards for Numeracy.....	251
Table 88: Grade level achieved by ALP girls in numeracy.....	252
Table 89: Tool Details.....	274
Table 90: Girls’ characteristics.....	276
Table 91: Potential barriers to learning and transition.....	277
Table 92: Output indicators .....	283
Table 93: Midline status of output indicators .....	284
Table 94: Output indicator issues .....	292
Table 95: Direct beneficiaries .....	293
Table 96: Other beneficiaries .....	294
Table 97: Target groups - by school.....	294

Table 98: Target groups - by age ..... 295  
Table 99: Target groups - by sub group ..... 295  
Table 100: Target groups - by school status ..... 296  
Table 105: Key analysis variables in household survey and learning assessments..... 298

## List of Figures

Figure 1: Learning scores for “panel” girls versus replaced and replacement girls.....	59
Figure 2: Distribution of learning scores by intervention versus comparison .....	75
Figure 3: Numeracy subtasks 1-6, by grade.....	86
Figure 4: Numeracy subtasks 7-11, by grade.....	87
Figure 5: Literacy subtasks 1-4, by grade .....	90
Figure 6: Literacy subtasks 5-8, by grade .....	91
Figure 7: Comparison of boys’ and girls’ literacy scores, by grade and intervention status .....	92
Figure 8: Comparison of boys’ and girls’ numeracy scores, by grade and intervention status .....	93
Figure 9: School Record Attendance as a Predictor of Learning .....	114
Figure 10: Caregiver Estimated Attendance as a Predictor of Learning .....	115
Figure 11: Transition pathways as a function of age, by intervention status.....	124
Figure 12: Count of enrolled girls per classroom, by grade level .....	126
Figure 13: Transition rates as a function of GEF participation, in varying samples .....	130
Figure 14: Share of teacher salaries paid by CEC, by round .....	148
Figure 15: Financial support of schools by CECs over the last year, according to parents.....	150
Figure 16: Parents reporting that their school has a functional CEC .....	154
Figure 17: Girls’ Attendance in Baseline and Midline – School Survey .....	167
Figure 18: Girls’ Attendance in Baseline and Midline – Household Survey .....	170
Figure 19: Girls’ Attendance by Grade – Household Survey.....	171
Figure 20: HH and SS Attendance Rates by HC Attendance Rate – Baseline.....	173
Figure 21: HH and SS Attendance Rates by HC Attendance Rate – Midline .....	174
Figure 22: School Survey Attendance by Teacher’s Use of Corporal Punishment.....	177
Figure 23: Household Survey Attendance by Teacher’s Use of Corporal Punishment .....	178
<i>Figure 24: Share of caregivers who report good or very good quality teaching over past 12 months .....</i>	<i>195</i>
<i>Figure 25: Index of active and participatory classrooms, by intervention status and round.....</i>	<i>198</i>
<i>Figure 26: Teachers’ use of formative assessments.....</i>	<i>199</i>
<i>Figure 27: Share of female students who believe girls and boys are treated differently at school .....</i>	<i>206</i>
<i>Figure 28: Frequency of classroom participation by boys and girls, by round .....</i>	<i>208</i>
<i>Figure 29: Teachers’ use of a harsh tone with boys and girls, by round.....</i>	<i>210</i>
Figure 30: Histogram of YLI Scores (in-school girls).....	215
Figure 31. YLI scores and learning outcomes .....	217
Figure 32. Midline girls’ responses (by intervention) .....	223
Figure 33: Midline girls’ responses (by intervention) .....	223
Figure 34: Index Score for Out-of-School Girls Over or Equal to 12 (Midline) .....	228
Figure 35: Index Score for In-School Girls Over or Equal to 12 (Midline) .....	229
Figure 36: Index Score for In-School Girls below 12 (Midline).....	229
Figure 37: Head teachers’ perceptions on likelihood of community support for girls’ vs boys’ school fees..	233
Figure 38: Reasons girl is not enrolled in school, according to caregivers .....	238
Figure 39: Reasons girl is not enrolled in school, according to caregivers, for ALP and OOS girls .....	239
Figure 40: Marriage and motherhood rates among ALP, in-school, and OOS girls.....	244
Figure 41: Distributions of learning scores for calibration girls.....	321

# Executive Summary

## Background

Despite ongoing efforts, learning outcomes in Somalia remain among the lowest in the region, particularly for girls. Boys and girls contend with different gender and social norms that tend to undermine their ability to stay in school, study and advance from grade to grade. Girls in Somalia are living in an environment undergoing deep transitions in social and gender norms, where traditional norms expecting women to primarily care for children in the home and assume responsibility for household tasks, and placing little value or emphasis on education for women coexist with new roles for women as entrepreneurs, heads of household and main breadwinners at home, thus increasing demand on girls' education. Since the time of the baseline, rural-rural migration has increased, predominantly as a result of economic hardship that has persisted among households that have been most heavily affected by drought. At the level of national government, MoE personnel tend to change frequently, leading to lack of continuity over time, but there is also increased funding for educational initiatives. It is in this context that CARE International launched SOMGEP and, following its successful completion, continued its programming through Somali Girls' Education Promotion Project – Transition (SOMGEP-T). The project, which began on May 1 2017 and is expected to close on October 31 2021, builds on evidence from SOMGEP and seeks to further address barriers and challenges Somali girls face related to attendance and learning outcomes. At proposal stage, the project was expected to reach a total of 27,146 marginalised girls; calculations based on up to date enrolment data indicate that the project is estimated to reach 27,722 in-school girls across 148 primary schools and 53 secondary schools in 22 target districts in Somaliland, Puntland, and Galmudug, as well as 5,140 out-of-school girls in the same locations.

SOMGEP-T aims to bring about sustainable improvements to the learning and transition outcomes of marginalised Somali girls. To address barriers and the causes of marginalisation, the SOMGEP-T Theory of Change (ToC) focuses on four key outputs: (1) Improved access to post-primary options, (2) Supportive school practices and conditions for marginalised girls, (3) Positive shifts on gender and social norms at community and individual girl level, and (4) Enhanced MoEs' capacity to deliver quality and relevant formal and informal education. Outputs are expected to contribute to the achievement of the project's four intermediate outcomes of attendance, retention, improved quality of teaching, and life skills development, which will in turn contribute to the long-term goals of improving learning outcomes, boosting transition rates, and ensuring the sustainability of changes brought about by the project.

The SOMGEP-T evaluation uses a mixed-methods, quasi-experimental design, involving a longitudinal panel of girls with a non-randomly assigned comparison group. The present study describes the results after four months of exposure to the intervention for in-school girls and presents the baseline findings for girls attending an alternative learning program (ALP). The midline sample comprises 63 schools, with 32 intervention schools and 31 comparison schools, plus 32 ALP sites (17 shared with the midline sample, 15 unique to the ALP sample). The primary findings from the evaluation are summarised below.

## Transition Outcome findings

At the time of the baseline, transition rates among cohort girls stood at just 50.8 percent among the intervention group. Due to the sampling design employed at the midline (excluding tracking of out-of-school girls, who will be tracked again in the 2019 evaluation round), midline transition rates cannot be compared to unadjusted baseline transition rates or to the targets set at that time. In comparable baseline and midline samples, transition rates rose from 74.7 percent to 81.5 percent. However, while transition rates in intervention areas increased by 5.9 percentage points, transition rates actually improved slightly more (1.8 points,  $p = 0.0.68$ ) in comparison areas, suggesting that improvements in transition are not currently being driven specifically, or exclusively, by project interventions. It is important to note, however, that reported transition rates capture only a subset of girls and do not account for areas – such as re-enrolment of out-of-school girls – where the project may be having an impact. At the same time, at least one specific project



intervention, participation in the Girls' Empowerment Forum, is associated with significant increases – of around 10 percentage points, depending on the particular sample studied and comparison made – in successful transition.

Successful transition among the largely in-school sample analysed in this report does not conform to the hypothesis that caregiver education, teaching quality or school infrastructure have a direct bearing on transition. But some factors consistently predict lower transition rates, including short-term economic deprivation and significant responsibilities at home. For instance, girls who face a particularly large burden of household chores are 11.7 points less likely to transition successfully. Notably, the largest overall gains from baseline to midline in terms of transition rates have occurred among girls with mental health disabilities, and girls who face a significant chore burden.

## **ALP Girls**

This evaluation served not just as the second evaluation round of SOMGEP-T's primary cohort of girls; it also served as a baseline assessment of girls enrolled in Alternative Learning Programs (ALPs). A sample of 365 ALP girls facilitated comparison to in-school and out-of-school cohort girls in terms of learning outcomes and a wide range of demographic, household, and personal characteristics.

As one might expect, learning scores among ALP girls consistently lagged behind in-school cohort girls of the same age. In other ways, however, ALP girls did not appear substantially different or more disadvantaged than in-school cohort girls. For instance, ALP girls were more likely to live in a household where the head of household had completed some formal schooling, and their caregivers were slightly more likely to be literate than those of cohort girls. Perhaps the single most important distinguishing characteristic of ALP girls was the number who were married, had been married, or who had given birth. Nearly one-quarter (24.7 percent) of ALP girls have been married, compared to just 0.8 percent of in-school girls. Even relative to out-of-school girls that are not enrolled in alternative education, ALP girls are more likely to be married and more likely to have given birth.

In addition, ALP girls often faced less tangible barriers than their household economic status or the educational background of their parents would suggest. ALP girls were more likely than in-school girls to have an extensive chore burden at home, and their caregivers had lower aspirations for their education. At the same time, ALP girls tend to have more supportive caregivers and lower chore burdens than out-of-school girls, occupying a kind of middle ground between in-school and out-of-school girls in many ways. An oversimplified portrait of ALP girls would suggest that they are much like other girls, but that their personal circumstances interrupted their schooling; rather than drop out entirely, enrolment in an ALP has allowed them to continue their schooling in a manner that is consistent with their circumstances and with the middling level of support they receive from their families.

## **Learning Outcome findings**

### *Literacy Outcomes*

For cohort girls, Somali literacy scores for the intervention group were 42.9 at the baseline and 59.8 at the midline (+16.9 percentage points from the baseline); scores for the comparison group were 40.3 at the baseline and 58.5 at the midline (+18.2 percentage points from the baseline).<sup>1</sup> Progress against the midline literacy target of 6.4 percentage points (over and above comparison) is -20 percent due to the fact that the

---

<sup>1</sup> Note, these figures include girls who were in grade 3 at the baseline and who are repeating grade 3 at the time of the midline. This set of girls who repeated grade 3 were later excluded in difference-in-differences estimation and the calculation of progress against midline targets. The determination was made to drop girls repeating grade-3 from the difference-in-differences analysis because their results cannot be easily entered into the tracking framework on the Outcomes Spreadsheet and because the difference in differences for their outcomes is extremely uncharacteristic of outcomes for cohort girls in grades 4-7 at the baseline, suggesting that these girls, or the classes that they attend, are qualitatively different from the rest of the sample in ways that are as yet unidentified.

amount of improvement in the comparison group was slightly greater than the amount of improvement in the intervention group (but not to a statistically significant degree).

### *Numeracy Outcomes*

Numeracy scores for the intervention group were 40.0 at the baseline and 51.9 at the midline (+11.9 percentage points from the baseline); scores for the comparison group were 38.0 at the baseline and 50.6 at the midline (+12.6 percentage points from the baseline).<sup>2</sup> Progress against the midline numeracy target of 6.5 percentage points (over and above comparison) is -12 percent due to the fact that the amount of improvement in the comparison group was slightly greater than the amount of improvement in the intervention group (but not to a statistically significant degree). In addition, it was determined that the numeracy boost training delivered to teachers during the first phase of implementation had not yet had a statistically significant effect on numeracy outcomes among the girls being taught by those teachers.

While the intervention as a whole did not have a detectable positive effect at baseline, it was found that girls who reported participating in the Girls' Empowerment Forum (GEF) had significantly higher literacy and numeracy scores than their peers who did not report participating in GEF. This finding suggests that at least some project interventions such as support for GEFs are beginning to have a measurable positive effect, even if the overall effects of the intervention are not yet evident.

Girls from pastoralist households were identified as vulnerable at the baseline, and they remain particularly disadvantaged at the midline, tending to have significantly lower literacy and numeracy outcomes than their peers. Qualitative data suggests that pastoralist households were hit the hardest by the effects of recent drought, leading to increased economic distress among pastoralist households that affected girls' abilities to have adequate study time and to attend school regularly. At the midline, disability status is not a predictor of significantly lower learning outcomes, with the exception of girls with vision disabilities who perform significantly lower than their peers. In terms of barriers that girls face, school and teaching quality were identified as problems at the baseline, and at midline poor principal and poor teacher performance remain two of the strongest predictors of lower learning outcomes. Despite a decreasing performance gap between boys' and girls' learning, unequal treatment of boys and girls by teachers in the classroom is strongly associated with lower learning outcomes.

### **Sustainability Outcome findings**

As measured at the midline, there has not been a measurable change in CEC financial support for schools that can be attributed to the project's impact. Qualitative reports indicate that CECs are taking a more active role in fundraising to support their schools, and CECs in general are supporting a larger share of teacher salaries. However, these improvements are occurring in both comparison and intervention schools, and CEC financial support is still relatively limited.

CEC functionality (as per parents' responses) has increased substantially since the baseline study, but these changes cannot be attributed to the intervention. Increases in school-level and community-level indicators of engagement have been larger for the comparison group than for the intervention group.

An increasing number of project schools are adhering to better-quality teaching practices, which is taken as a proxy for adherence to a broader set of implementation standards. In particular, the number of teachers that report using formative assessments has increased from 43.1 percent at baseline to 71.9 percent at the midline. It is important to note that this measure is only a proxy for a broader set of standards; moreover, future evaluation waves – to the extent that they wish to focus on the use of formative assessments – should consider alternative measurements that might combat social desirability bias, including requesting additional evidence of their use.

---

<sup>2</sup> As above, these figures include girls who were in grade 3 at the baseline and repeated grade 3 at the midline.

System-level sustainability measures are the most difficult and tenuous to measure, but the project has shown progress on this front as well. Ministry of Education officials appear to be aware of ALPs and vocational training as viable options for girls, and officials also confirmed an increased focus at the REO level on girls' education, through the establishment of dedicated gender units.

### **Marginalisation Analysis and Gender Analysis (including GBV)**

Girls belonging to pastoralist households were the most marginalized at baseline and remain at highest risk at the midline. They tend to have lower learning outcomes, lower attendance rates, and poorer transition outcomes. The qualitative data provides less evidence of negative stereotypes against pastoralist girls at the midline, but the drought has maximised the challenges for pastoralist households by threatening their livelihoods and erecting new financial barriers to keeping girls in school by creating incentives to take them out of school to help with income-generating activities or otherwise helping out in an economically distressed household. There is still a gap between girls' and boys' numeracy and literacy outcomes, but girls show signs of closing the gap, and grade 3 (4) girls are performing at the same level as grade 3 (4) boys.<sup>3</sup> At the midline, foundational skill gaps have shifted as learners have advanced grades and learned more, but boys and girls still demonstrate the same fundamental skill gaps. In terms of attendance, there are no significant differences between boys and girls.

The results of the new risk mapping tool deployed in the midline suggest that girls feel safest and happiest at school and in their homes and that they feel least safe on roads/pathways. Girls feel particularly unsafe on roads at night, with many citing the presence of gangs and thieves, as well as fears of being raped, kidnapped, or killed. Despite the fact that girls most commonly mentioned feeling safe at school, girls do not feel safe at times in certain areas of their schools where they face harassment from boys. One of the most notable findings from the new risk mapping exercise was that girls face verbal, physical, and sexual abuse from boys both in school and in their communities. Boys reportedly throw stones at girls, say inappropriate things to them, and push toilet doors open. Equally notable is that girls with disabilities face significant verbal and physical abuse, and in fact, a number of girls interviewed for the midline reported having become disabled after being physically attacked. Although girls with disabilities report having friends in school and support networks at home, they also report being teased by other students for their disabilities and getting into fights in school.

### **Girls with Disabilities**

The midline evaluation revealed a higher proportion of girls in the panel data with mental health impairments than in the baseline in both intervention areas (14 percent vs 5.5 percent) and comparison areas (13.8 percent vs 5.2 percent). However, this increase is mainly due to the broader midline sample of primary caregivers who were asked questions about their cohort girls' mental health. In the baseline, only primary caregivers whose girls were under 12 years of age were asked questions about mental health, whereas the midline evaluation was designed to ask primary caregivers of cohort girls of all ages mental health questions after a relatively high proportion of girls with mental health impairments was observed in the baseline. All other types of impairments, visual, hearing, mobility, cognitive, self-care, and communication impairments, which were rare in the baseline, did not substantially change in the midline.

Re-contact rates were within the expected range for the project, which anticipated and planned for high panel attrition. The most common reason a girl needed to be replaced – nearly half of all cases – was because her household had moved out of the community. Re-contacted girls differed from girls who could not be re-contacted and necessitated replacement primarily in terms of their age and geographic region. Seasonal migration and low household economic status did not predict higher attrition rates, surprisingly.

---

<sup>3</sup> Throughout the report, the following syntax is used to refer to grade level: the baseline grade is stated in the text followed by the midline grade in parentheses. For example, "grade 3 (4)" means that the baseline grade is grade 3 and the midline grade is grade 4.

## Intermediate Outcomes findings

### *Attendance*

In the headcount survey, the intervention did not appear to have a significant impact on the attendance rates across the girls' and boys' group. While insignificant, the average headcount attendance rates have marginally improved from the baseline to midline in the intervention groups for both boys and girls whereas the yesterday's and today's average attendance rates have decreased very slightly over time. In particular, the boy's average headcount rate in the intervention group had more improvements from the baseline (80.9 percent) to midline (82.91) compared to the boys' rate in the comparison group (82.79 and 79.09, respectively). Similarly, girls' average headcount increased from 81.8 percent in the baseline to 82.0 in the midline, although their counterpart's average headcount rate decreased from 84.0 in the baseline to 80.9 in the midline.

The school survey attendance rates for intervention areas in the midline is 91.7 percent, a marginal decrease from the baseline 91.9 percent attendance rate. Comparison schools experienced a similar drop in the school survey attendance rates, from 94.5 percent in the baseline to 94.3 percent in the midline. A significant positive intervention effect is not observed in difference in difference modeling.

However, there was a significant improvement from baseline to midline in attendance recordkeeping that was captured by the school survey. Only 49.1 percent of cohort girls had enrolment records in the baseline while 70.3 percent of cohort girls had enrolment records in the midline. The intervention effect is positive, but not significant. In addition, to more enrolment records being kept in the midline, it seems as if the accuracy of the recordkeeping has also improved as the relationship between the school survey attendance rates and the headcount survey attendance rates is positive and significant.

### *Life skills*

The girls' average YLI score has marginally improved in the intervention group from the baseline to midline, but this improvement has been less than their counterparts' average YLI score in the comparison group. However, girls in the intervention group scored higher, on average, in the life skills (except for the younger out-of-school girls who were not included in the analysis) assessments compared to the girls in the comparison group. In particular, girls in the intervention group feel less nervous when speaking in front of an adult and in a group of people as well as doing math and answering questions in class. Yet, none of these differences between the intervention and comparison group were found to be statistically significant. The internal consistency level is high for the YLI index and reasonable for the life skills index. As a result of participating in the SOMGEP-T project, girls with disability seem to be the only group to have significant improvement in the level of their confidence speaking in front of a group.

The qualitative evidence collected from the FGDs with mothers and teachers and KIs with girls with disabilities also suggest that girls' engagement and participation is improving generally, despite some barriers that girls face. Although shyness is one of the factors preventing girls from having effective participation, the girls themselves state that they feel confident speaking in their homes, school and communities.

### *School Management*

The CEC plays a crucial role in the management and governance of schools. From the head teachers' perspective, it was found that the SOMGEP-T project did not have a significant impact on CEC activities (such as monitoring teachers' attendance, facilities, teaching quality, or students' attendance and drop out during their last visit) in the intervention as a result of their participation. The qualitative data showed that head teachers have negative perceptions toward the work CECs are doing, which partially may be stemming from the tensions that exist between CEC members and head teachers.

Yet, the primary caregivers in the intervention group generally have a more positive view about the role of CECs in their children's schools. Except for the students in the intervention group who tend to receive more

financial support from CECs from a baseline of 9.5 percent to a midline of 16 percent, none of the difference-in-differences between the intervention and comparison groups and across subgroups were significant in terms of how well the CECs are functioning or what activities they are carrying out.

### *Teaching Quality*

On almost every metric employed, teaching quality has improved from baseline to midline. However, as with many of the other outcomes assessed in this report, the rate of improvement is nearly indistinguishable between intervention and comparison schools. For instance, while the share of teachers observed asking students open-ended questions increased 16 percentage points in intervention schools, comparison schools registered a 26 point increase. Where intervention schools stood out was in the expanded use of formative assessments: the share of intervention school teachers who report their use increased from 43.1 to 71.9 percent; no comparable increase was found in control schools.

### *Recommendations*

The following recommendations follow from the report's findings:

- The findings related to learning suggest that teaching quality remains one of the most critical determinants of learning and should be monitored closely going forward. However, it should not be concluded that teaching quality is not responding to project interventions; rather, it is most likely the case that teachers have not yet had time to demonstrate improvement on the basis of project interventions.
- Because conflict-related school closures appear to be the primary cause of attenuated learning among girls in some locations, it would seem that intensified support for supplemental learning programs will be the most appropriate way of enabling girls who have missed significant amounts of school to catch up with their peers.
- Girls belonging to pastoralist households remain at high risk of missing days of school, and the economic distress in their households may reduce their study time and even their food security in ways that significantly affect their ability to learn in school. In particular, pastoralist girls who are re-enrolling in formal education will require additional tutoring or help with their studies in order to compensate for the fact that many have been out of school or have had lower-than-average attendance levels.
- The collection of attendance and enrolment records has improved dramatically from baseline to midline, however substantial gaps remain. The effort to further improve attendance can be supported by providing training and resources that will help teachers maintain attendance records. Furthermore, additional interventions should help teachers, principals, and other stakeholders use attendance record to identify and target interventions to at-risk girls.

## 2. Background to project

### 2.1. Project overview

The Somali Girls' Education Promotion Project – Transition (SOMGEP-T) aims to address the barriers that Somali girls face in regularly attending school and acquiring literacy, numeracy, English and financial literacy skills and to create the conditions for successful transition of in-school girls into new education levels and of out of school girls into school, alternative education or life skills training. The presence of sustainable mechanisms and social norm change to maintain and expand gains will also be assessed as a third outcome (sustainability). SOMGEP-T began on May 1 2017, and will close on October 31 2021. The project will cover 148 primary schools, 53 secondary schools, and the respective catchment areas. SOMGEP-T will be implemented in 22 target districts in Somaliland, Puntland, and Galmudug.

At the time of its design, SOMGEP-T was expected to reach a total of 27,146 marginalised girls, of whom 16,863 were enrolled in school at the Somali Girls' Education Promotion Project's (SOMGEP) endline; 1,583 were out of school and living in villages targeted by SOMGEP; and 7,834 new entrants, who will benefit at no extra cost. Only schools and communities who benefitted from the initial SOMGEP intervention will be tracked by SOMGEP-T, therefore determining the effects of the GEC investment in the target areas through time.

### 2.2. Project context

Following the ousting of President Siad Barre's military regime in 1991, Somalia's central government collapsed, and the country descended into a civil war as numerous actors—including clan warlords, pirates, radical groups, and others—vied for power. Despite the efforts of the Somali people and outside forces from three separate international peacekeeping missions (UNOSOM I, UNITAF, and UNISOM II)<sup>4</sup> and one regional peacekeeping mission which remains active today (the African Union Mission in Somalia, or AMISOM), the civil war persisted for more than two decades. The establishment of a transitional government in 2004 marked the first major step toward peace. In 2012, Somalia held its first presidential election since 1967 and swore in its first formal parliament in more than 20 years. In the same year, the National Constitutional Assembly adopted The Provisional Constitution of the Federal Republic of Somalia (FRS), officially establishing the Federal Government of Somalia (FGS). The electoral process of 2016 marked the first successful transition of power in Somalia since the creation of the FGS.

The people of Somaliland, Puntland, and Galmudug are involved in various economic activities, but livestock is the lifeblood of the regions' economy. The livestock sector constitutes up to 60 percent of the Somaliland government's revenue base and 20 percent of the country's GDP. Furthermore, half of Somaliland's population are agropastoralists who produce crop (sorghum and maize) and engage in livestock rearing. Puntland similarly has large livestock sector but also has the capacity to develop its leather tanning and production sector. Galmudug has the potential for a strong livestock exportation and also owns untapped natural resources, such as meerschaum, fluorspar, and uranium.<sup>5</sup>

These zones have experienced widespread drought in the last decades with devastating impact on communities and their livelihoods. Drought has had the highest estimates of damages in Puntland (\$934

---

<sup>4</sup> The United Nations Operation in Somalia I (UNOSOM I), the Unified Task Force (UNITAF), and the United Nations Operation in Somalia II (UNOSOM II). UNOSOM I and II, authorized by Resolutions 751 and 814, respectively, were UN-led peacekeeping missions. UNITAF was a Coalition of the Willing, led by the United States but joined by more than 15 other governments and their forces.

<sup>5</sup> UNICEF. (2017). UNICEF Somalia Situation Report. Retrieved from reliefweb: [https://reliefweb.int/sites/reliefweb.int/files/resources/UNICEF%20Somalia%20Humanitarian%20Situation%20Report%20%231%20-%20February%202017\\_0.pdf](https://reliefweb.int/sites/reliefweb.int/files/resources/UNICEF%20Somalia%20Humanitarian%20Situation%20Report%20%231%20-%20February%202017_0.pdf).



million), followed by Somaliland (\$874 million) and Galmudug (\$395 million). The damage inflicted by the drought varied across sectors. For example, the productive sector, including irrigated and rain-fed crops, livestock, and fisheries, accounts for 65 percent of damages and losses in Somaliland, 62.7 percent in Puntland and 60.1 percent in Galmudug (United Nations et al., 2018). Somalia has historically received low and inconsistent rainfall. During September 2016 to June 2017, for example, 30 out of 42 districts in central and southern Somalia and 26 out of 32 districts in Puntland and Somaliland received significantly below-average rainfall than the previous year (United Nations et al., 2018).

Despite these successes, the effects of the civil war are still evident today, and Somalia's stability continues to be threatened by violence, poor governance, uneven development, and humanitarian crises. Within the education sector, the post-war years have been marked by slow growth. With the complete collapse of the tax administration system, domestic revenue (taxes plus fees) represents just 2.8 percent of GDP, making it difficult for the government to provide services.<sup>6</sup> In 2016, education and health accounted for only 2.5% of the budget, and although US \$4.7 million was approved for education, only US \$1.8 million was executed.<sup>7</sup> In the central and southern regions alone, more than 75% of the public schools that existed prior to the civil war were destroyed or closed,<sup>8</sup> and across the entire country, an estimated 90% of schools were destroyed.<sup>9</sup>

The education sector has been hard hit by drought, famine, and economic hardship, which have led some children to drop out of school. While figures from the government may be understated, UNICEF estimates that the highest proportion of children dropping out of school as a result of the drought would be in Puntland (10.2 percent of enrolled children), followed by south central (8.3 percent) and then Somaliland (3.3 percent).<sup>10</sup> Displacement has affected the country's educational infrastructure with some schools being abandoned and others becoming overpopulated where internally displaced people have settled in the host communities (United Nations et al., 2018). Currently, thousands of people are internally displaced including 870,000 in South and Central Somalia, 130,000 in Puntland, and 40,000-80,000 people in Somaliland.<sup>11</sup> Moreover, lack of water, sanitation and hygiene (WASH) services has affected the lives of millions of people in all three zones. There is an estimated of 1.1 million people in urgent need of WASH assistance in Somaliland and Puntland and 1 million people in Galmudug.<sup>12</sup>

In the absence of adequate public education options, various stakeholders, including NGOs, private institutions, religious groups, and others, have attempted to fill the gap. These groups have made progress in rebuilding infrastructure and providing much needed financial assistance but have also introduced a new set of challenges to the education system. The institutions established by these groups often have their own curricula and examinations, operate on fee systems, and do not have uniform criteria for determining fee exemption.

---

<sup>6</sup> World Bank Group (2017). Somalia Economic Update – Mobilizing Domestic Revenue to Rebuild Somalia.

<sup>7</sup> Ibid.

<sup>8</sup> UNICEF (2017). UNICEF Somalia Quarterly Education Bulletin - July 2017.

<sup>9</sup> Barakat, Connolly, Hardman, Lewis, Lineker, Menkhaus, Rzeszut and Shanks, 2014. Beyond Fragility – a conflict and education analysis of the Somali context.

<sup>10</sup> United Nations, World Bank, European Union, Government of Somalia, & GFDRR. (2018). Somalia Drought Impact and Needs Assessment (Vol. 2): Sector Report. Washington, D.C: UNOCHA. Retrieved from reliefweb: <https://reliefweb.int/sites/reliefweb.int/files/resources/122991-v2-Revised-GSURR-Somalia-DINA-Report-Volume-II-180111-Digital.pdf>.

<sup>11</sup> Drumtra, J. (2014). Internal Displacement in Somalia. Washington DC.: Brookings Institution. Retrieved from Brookings Institution: <https://www.brookings.edu/wp-content/uploads/2016/06/Brookings-IDP-Study-Somalia-December-2014.pdf>.

<sup>12</sup> UNICEF. (2017). UNICEF Somalia Situation Report. Retrieved from reliefweb: [https://reliefweb.int/sites/reliefweb.int/files/resources/UNICEF%20Somalia%20Humanitarian%20Situation%20Report%20%231%20-%20February%202017\\_0.pdf](https://reliefweb.int/sites/reliefweb.int/files/resources/UNICEF%20Somalia%20Humanitarian%20Situation%20Report%20%231%20-%20February%202017_0.pdf)

Within the current public education system, the quality, availability, and outcomes of education vary drastically by area. However, the public education systems in all areas follow the same general structure. Each has its own Ministry of Education and Higher Education (MEHE), with offices at the regional and district level for planning and coordination purposes. In Somalia/Somaliland, early childhood education (ECE) is offered to children aged 3-5 through formal pre-primary schools or Quranic schools. Children in Somalia/Somaliland then enter their primary education, which they attend from the age of 6 to the age of 13 through formal primary schools, Integrated Quranic Schools (IQS), and in the case of older students who have not previously attended school or dropped out in early primary, through Alternative Basic Education (ABE). Late enrolment is common: SOMGEP's baseline (2014) found that only 31% of the girls and 47% of the boys are enrolled at age 6,<sup>13</sup> largely due to the practice of prioritizing Quranic education for young children. There are indications that Quranic school attendance supports students to learn to decode in Arabic and build foundational literacy skills; these skills later facilitate the acquisition of reading skills in Somali upon entering formal education.<sup>14</sup> Following completion of their primary education, students enter their secondary education, intended for children aged 14-17, or Technical Vocational Education and Training (TVET), which can last for the same duration of time as secondary education or longer. After completing secondary school, students can continue on to the university level, which is intended for children aged 18-21, or can choose to continue their studies through TVET. Pre-service teacher training is also an option for secondary school graduates.

As a result of the civil war and the lack of coordination across the education sector, inequities in education are present across genders, groups, and regions. In 2014, UNFPA published the results of its Population Estimation Survey, the first extensive household survey to be carried out in Somalia since 1975, when the government last published census results.<sup>15</sup> The survey finds that urban areas have the best adult literacy rates, the highest level of education completed for the out-of-school population, and the highest school enrolment of persons currently in school. In fact, enrolment in urban areas is approximately double that of enrolment in rural areas, and triple that of nomadic populations. The survey also finds that there is a positive relationship between wealth status and school enrolment.

Across education indicators, males have fared better than females. Enrolment rates are slightly higher for males than females, and there is an 8 percentage point gap in adult literacy in favour of males. The gender parity index (GPI) rating worsens in higher levels of education, dropping from 0.986 at the primary level to 0.916 at the secondary level to 0.688 at the tertiary level. Social gender norms around the roles of women in society are responsible for creating barriers to girls' enrolment, retention, and school performance. Whereas men are expected to pursue employment outside the home, women have traditionally been expected to care for the children and assume responsibility for household tasks. Data from SOMGEP's midline and endline assessments provide evidence that these norms are changing as an increasing number of women are now taking new roles in society, largely as a result of migration and conflict dramatically raising the proportion of female heads of household, but a variety of factors still limit girls' interest in school and colour adults' perceptions of the importance of their education. Early marriage leads to drop-out or prevents girls from ever attending school. In addition to seasonal migration, which affects both girls and boys, girls face a number of unique challenges that lead to absenteeism. Girls are often recruited by their mothers to assist in household tasks, which causes them to miss school or enroll late. They also tend to withdraw from school during their menstrual periods, which can be particularly painful for girls who have undergone Type III female genital mutilation (FGM). Historic low rates of access to education for women, traditional gender norms which limit women's mobility and the ability to work outside the home, particularly for those who are mothers, make the

---

<sup>13</sup> CARE (2014) Somali Girls' Education Promotion Project – Baseline Study Report

<sup>14</sup> SOMGEP's evaluation studies have consistently noted an accelerated acquisition of reading skills, including reading comprehension, in early primary. Quranic schools in Somalia prioritize actual decoding and reading of Arabic, as opposed to simple memorization of the Holy Quran; it is therefore hypothesized that the early experience in decoding Arabic may have a positive effect on decoding the Latin alphabet used for Somali.

<sup>15</sup> UNFPA (2016). Population Composition and Demographic Characteristics of the Somali People.



low recruitment of female teachers a particular problem; as a result girls lack role models, and the proper support and counselling services at school that might encourage them to continue their education, particularly upon reaching adolescence.

The composition and characteristics of the Somali population underscore the importance of these findings. Although there has been a decrease in nomadic populations, currently approximately 42 percent of the Somali population lives in urban areas, 26 percent in nomadic settlements, 23 percent in rural settlements, and nine percent in IDP camps. Among the general population, the majority of households in all areas are headed by men, with the largest proportions of male-headed households found in nomadic and rural communities (92.9% and 81.3%, respectively) and the lowest in urban areas and IDP camps (77.6% and 75.6%, respectively). An analysis of the patterns observed in SOMGEP's evaluation studies suggests, however, that the number of female-headed households is consistently increasing and exceeding that of male-headed HHs in the rural and remote areas where the project operates, potentially as a result of migration, displacement and conflict. 40% of the households surveyed for SOMGEP's midline were female-headed, compared to 43% at the endline.<sup>16</sup>

The Singulate Mean Age at Marriage (SMAM)<sup>17</sup> is 24.7 years for males and 23.1 years for females. Males and females from rural areas are the most likely to marry young, whereas those from urban areas are the most likely to delay marriage. Among females, those who have completed tertiary education have the highest SMAM. Of the 58% of individuals in the population who are currently married, 72.2% have not completed any level of education. The prevalence of child marriage in Somalia is estimated as 45%<sup>18</sup>, although findings from SOMGEP-T baseline indicate that the actual prevalence in the project's target areas may be considerably lower. Cultural perceptions often limit the recognition of child marriage as an issue; qualitative data suggests that many communities traditionally acknowledge individuals age 15 and above as adults, not children.

Households living in rural areas are heavily affected by the effects of climate change. The prolonged drought experienced since 2015 resulted in massive loss of livestock and large levels of displacement. The loss of livelihoods is exacerbating malnutrition, with 1.4 million children estimated to be currently affected.<sup>19</sup> As a result of the drought, 948,500 people<sup>20</sup> are estimated to have been displaced since late 2016, with large numbers moving into cities, placing additional pressure on an already fragile education system. As of January 2018, 866,000 people are affected directly by the emergency situation and poor harvests are expected later this year.<sup>21</sup> The vulnerability to the negative effects of climate change (such as threats to livestock livelihoods and therefore household income) and the use of migration as a coping mechanism are likely to continue to affect the economic and the social landscape of the country and will be a key underlying factor shaping the education outcomes of the children and adolescents targeted by this program.

It is important to note that while Somalia does face major challenges related to education, a number of factors are leading to rapid shifts in the country. For example, Somalia's diaspora, estimated at 1.5 million people or about 15% of the total population, are playing a role in the country that is disproportionate to their numbers.<sup>22</sup> Diaspora not only support education, but also are the main source of private capital and investment. Foreign investment from diaspora serves as an important source of funding for both the

---

<sup>16</sup> JBS International (2016) Somali Girls' Education Promotion Program – GEC Endline Evaluation Report

<sup>17</sup> Singulate Mean Age at Marriage refers to the average number of years of single life before age 50 of the population born in the same year.

<sup>18</sup> UNFPA (2012) *Marrying Too Young: End Child Marriage*.

<sup>19</sup> According to UNICEF's May 2017 Humanitarian Response Bulletin.

<sup>20</sup> Famine Early Warning Systems Network - <http://www.fews.net/east-africa/somalia/food-security-outlook/october-2017>.

<sup>21</sup> Famine Early Warning Systems Network (<http://www.fews.net/east-africa/somalia>)

<sup>22</sup> Barakat, Connolly, Hardman, Lewis, Lineker, Menkhous, Rzeszut and Shanks, 2014. *Beyond Fragility – a conflict and education analysis of the Somali context*.

education sector and the private sector, and communities across the project's locations have become dependent on remittances in order to access basic privatized services. For example, it is estimated that total private transfers from Somali diaspora are the third largest contributor to Puntland's GDP.<sup>23</sup> Diaspora and others operating in the private sector (businesses, NGOs, faith-based groups) now play a significant role in providing education across the project's three zones. Another important factor is mobile penetration. Currently, smartphones account for less than 25% of connections in Somalia, but smartphone adoption is expected to double by 2020 to reach 45%.<sup>24</sup> Lastly, the roles of women are changing—evidence from SOMGEP suggests that there are more women leading households, and it is estimated that over 60% of the owners of small businesses are women. Given the historically low levels of education, women are unequipped for these new roles, but these shifts will have important implications for future development programming in the region.

SOMGEP-T has purposefully targeted the poorest and most excluded locations in Somaliland, Puntland, and Galmudug. The project will target the most marginalized portions of the population in these areas, which depend on pastoralism and have suffered from the repeated occurrence of long droughts. SOMGEP-T's implementation area includes large proportions of villages that are facing widespread emergency (IPC phase 4) and high risk of famine as a result of the droughts. Additionally, ninety-nine percent of the girls sampled in SOMGEP's baseline were marginalized, and the project's endline indicated that there were even further declines in indicators related to household conditions and the ability to invest in education, and that there was a sharp increase in displacement due to the drought.

## 2.3. Project Theory of Change and assumptions

### Theory of Change Overview

The long-term goal of SOMGEP-T is to bring about sustainable improvements to the learning and transition outcomes of marginalised Somali girls. Marginalised girls who are targeted under SOMGEP-T are expected to exhibit statistically significant improvements in learning outcomes (literacy, numeracy, and financial literacy) and transition outcomes (transition rate) as compared to a comparison group; targeted schools, communities and government institutions are expected to demonstrate indications of sustainability. To achieve its long-term outcomes and create a more supportive environment for girls, the project will focus on addressing the underlying causes of marginalisation through influencing stakeholder attitudes and promoting social change at the household, school, community and policy/governance levels.

SOMGEP-T defines marginalised girls as those who face the intersection of multiple barriers to access education and once enrolled, to remain in school after Grade 3.<sup>25</sup> SOMGEP's studies have identified that the barriers marginalised girls face include extreme poverty, pastoralism, displacement, being over age for their grade, a high degree of exposure to violence/ conflict, orphan status, disability, belonging to a minority clan, and having an illiterate mother (who is often experiencing financial hardship as a female head of household).

More generally, barriers to girls' education in Somalia can be categorized as demand-side barriers and supply-side barriers. Demand-side barriers include traditional gender and social norms (early marriage, chores, girls' low agency, gender-based violence (GBV)), poverty and high vulnerability to the negative effects of climate change, high absenteeism (seasonal migration, chores), perceptions of disconnect between education and the local market, and armed conflict. Supply-side barriers include limited provision of secondary education and poor infrastructure, limited number of qualified teachers, low teacher capacity to teach higher numeracy skills and English as a second language, lack of catch-up opportunities/ remedial

---

<sup>23</sup> Ibid.

<sup>24</sup> GSMA (2018). GSMA Mobile Economy.

<sup>25</sup> CARE SOMGEP-T MELF Final

education for pastoralist children, and limited capacity of school leadership and education officials to address absenteeism, dropout and poor learning outcomes.

In addressing the barriers to girls' education in Somalia, SOMGEP-T will focus on four key domains of change, or direct outputs: (1) improving access to post-primary options; (2) fostering supportive school practices and conditions for marginalised girls. (3) promoting positive shifts on gender and social norms; (4) enhancing the capacity of MoEs to deliver quality education. According to SOMGEP-T's ToC, if CARE International and its partners focus on these domains of change, then the number of girls who access, receive, and complete a quality primary and secondary education will increase.

## Expected Outcomes

*Long-term outcomes:*

- (1) Learning: The number of marginalised girls supported by GEC with improved learning outcomes (literacy, numeracy, financial literacy).
- (2) Transition: The number of marginalised girls who have transitioned through key stages of education, training, or employment.
- (3) Sustainability: The changes brought about through the project which increase learning and transition through education cycles are sustainable at the community, school, and system levels.

*Intermediate Outcomes:*

- (1) Attendance
- (2) School governance
- (3) Improved quality of teaching
- (4) Life skills development

## Outputs and Key Activities

Through its key activities, SOMGEP-T will deliver four key outputs to improve the learning and transition outcomes of marginalised Somali girls and empower them to engage in the local economy and decision-making processes in the future. SOMGEP-T's outputs and activities are outlined below.

*Output 1: Improved access to post-primary options*

Girls will be supported to transition into either formal secondary schools through grants (bursaries)<sup>26</sup> for poor families, provision of an ALP developed in partnership with the Ministries of Education (MoEs) and communities, and development and strengthening of community education committees (CECs).

Key Activities:

- Work with MoE to develop and implement ALP
- Develop girls' life skills in upper primary through ALP, including leadership skills, financial literacy and business selection and management of income generation activities
- Develop CECs to improve retention and transition
- Provide partial grants to girls from poor families
- Equip two boarding schools for girls with furniture / learning materials and promote girls' enrolment

GESI Transformative Adaptations, Gender:

- Prioritise the recruitment of female teachers for ALP.
- Advocate for teachers training colleges to have a quota for female graduates.

---

<sup>26</sup> School fees are paid directly to schools; in addition, girls receiving grants also receive uniforms and basic support items from the projects

- Lobby for inclusion of women in CECs structure and encourage effective participation
- Work with CECs to promote participatory and inclusive review of SIP's and ensure that the SIP clearly outline feasible and relevant initiatives for promoting girls education

#### GESI Transformative Adaptations, Social Inclusion:

- Mapping the spread of nomadic and pastoralists households and develop appropriate strategies for reaching these social excluded sub populations
- CEC and host communities intensify efforts to enrol girls from nomadic households as well those from pastoralist families

#### *Output 2: Supportive school practices and conditions for marginalised girls*

The project will boost numeracy outcomes and English skills among primary and secondary students, providing remedial support to struggling students as well as those with high absenteeism rates, particularly pastoralist girls, and supporting the school leadership to track attendance, learning, retention and transitions, therefore increasing the chances of marginalised girls building foundational skills, completing primary school and succeeding in secondary education.

#### Key Activities:

- Train teachers on improved delivery of literacy and English language, supported by digital content in all 148 primary and 53 secondary schools
- Train teachers on improved delivery of numeracy in all 148 primary and 53 secondary schools
- Train teachers to provide structured remedial support to students at primary and secondary level
- Train and coach teachers to deliver the ALP curriculum
- Construct additional classrooms in remote primary schools; build water facilities in new secondary schools; and provide solar chargers for mobile devices/ tablets and sanitary pads to schools
- Incorporate life skills and financial literacy training into Girl's Empowerment Forums (GEFs) and Boys' Empowerment Forums (BEFs)
- Provide career guidance in schools

#### GESI Transformative Adaptations, Gender:

- Review of inactive empowerment forums and revitalises non-functional empowerment forums
- Identify women role models in private sector to participate in school-based career guidance sessions
- Security risks as a result of lack of perimeter wall the project need to explore options with CECs, especially use of durable, less costly, locally available and environmentally friendly materials

#### GESI Transformative Adaptations, Social Inclusion:

- Remedial sessions to be rolled out at the same time in a number of schools preferably starting with remote schools to maximise exposure. This approach will ensure that remote schools have maximise exposure.
- Schools with few and less qualified teachers will be prioritised in all trainings.

#### *Output 3: Positive shifts on gender and social norms at community and individual girl level*

Through promoting positive shifts on gender and social norms, the project will create an environment where girls and boys are equally supported to attend school, their skills are valued, there are higher expectations for their achievement, and where girls and boys are safe from harmful practices. It is expected that these activities will not only encourage parents to send their girls to school but will also encourage girls to stay in school by creating a safe environment for them and emphasizing the importance of education in relation to other social pressures that typically cause girls to drop out, such as marriage.

#### Key Activities:

- Engage community-level stakeholders including religious leaders, women's groups, men and boys
- Expand and strengthen GEFs and create BEFs to develop leadership and mentorship skills
- Provide adult literacy and financial literacy classes for mothers
- Support the financial empowerment of mothers through savings groups (VSLA), business selection, and business coaching and mentoring

#### GESI Transformative Adaptations, Gender:

- Greater inclusion of men as agency of change to spearhead the mobilisation of the community to promote girls education

#### GESI Transformative Adaptations, Social Inclusion:

- Community awareness to reach various subgroups including nomadic and pastoralists, rather than focusing on host communities

#### *Output 4: Enhanced MoEs' capacity to deliver quality and relevant formal and informal education*

MoEs' staff, local education officers will be supported to develop robust governance and support structures, taking an active role in improving girls' retention and transition rates, overseeing the implementation of quality standards and data management systems, and identifying and addressing barriers to learning using a gendered lens. MoEs are uniquely positioned to send a strong, positive message about the importance of girls' education.

#### Key Activities:

- Strengthen Gender Departments capacity to improve girls' education outcomes through trainings, development of action planning and provision of incentives to retain the gender focal points especially in rural areas
- Support quality assurance and standards (QAS) functions at all MoE levels
- Provide support to Regional Education Officers (REOs) and District Education Officers (DEOs) to mainstream improved teaching practices and address retention/ transition issues
- Work closely with MoE on non-formal education (NFE) for mothers and entrepreneurship skills for girls
- Development of project IEC materials in conjunction with MoE for use at stakeholder advocacy and promotion events

#### GESI Transformative Adaptations, Gender:

- Include the recruitment of female, teachers, head teachers, Regional Education Officers (REOs) and District Education Officers (DEOs) in the advocacy activities.

#### GESI Transformative Adaptations, Social Inclusion:

- Support Regional Education Officers (REOs) and District Education Officers (DEOs) to access remote schools, which are often neglected because lack of transport by incorporating school visits by the REOs during project implementation

In addition to these initially planned activities and GESI-relevant adaptations, the following general adaptations were made by the project in response to baseline findings:

#### Adaptation 1, to improve learning especially in numeracy:

- Extensive coaching sessions for teachers in struggling schools (from April to May) before schools close using the numeracy Level 1 module developed in Phase One for basic number operations and coaching guidelines developed recently. This will help build teachers' delivery of the content and students' foundation skills in maths.
- Subject to availability of funds the project is exploring how to achieve maximum exposure by running a remedial learning sessions during the June to August school break. Funds will be particularly required to commit teachers on monthly performance based contract on number of lessons delivered

during the 3 months. The remedial support will be guided by grade specific tasks identified through the baseline process.

- The project will also consider inclusion of participatory and interactive basic maths games in GEF and BEF manuals. The original plan was to only focus on life skills and include aspects of financial literacy in the GEF manuals. This will be borrowed from the Numeracy and ALP modules and will not involve extra cost.

Adaptation 2, in response to low baseline achievement levels in literacy:

- Trimming the English literacy test (removing the upper reading comprehension task and the written tasks),
- Addressing limited reading fluency and vocabulary, matching the electronic platform with a clear messaging to the teachers on its use (when, which dosage, with whom), as many of the students are not learning it in school at all.
- Exploring positive deviants (the five schools that performed very well, particularly Ilays), to learn what are the successful strategies there.

Adaptation 3, involves the project adopting a stronger (and more specific) monitoring of CECs which encompasses the periodic assessment of CECs functionality and fidelity of implementation. In order to provide additional support to CECs, adaptations will be tailored to respond to the specific issues arising from the assessments. The identified issues will be addressed during the coaching of CECs. Additionally, the project intended to:

- Further query data to see where the problem is in relation to the functionality of CECs.
- Develop a coaching guide for CECs to be reviewed and include more details for staff/MOE.

Adaptation 4, in order to improve GEF impact, includes:

- Tracking GEF activities and verifying their functionality, as there is a likelihood that many aren't meeting/ functional after the pioneer group of older girl's transition to higher levels of education.
- For functional groups, use self-monitoring/ reflection/storytelling tools whenever possible, as opposed to more traditional checklists.
- Non-functional groups will require mobilization and refresher trainings, identifying and addressing reasons for disbanding/ lack of functionality.
- Focus the Gender Department Annual work plan on supporting the GEFs as a matter of priority.

Adaptation 5, in order to address high proportion of teachers observed using corporal punishment at the baseline, adaptations include:

- Including coaching that models other, non-violent ways of disciplining children.
- Mainstreaming alternative, non-corporal discipline strategies throughout project activities: i.e. Teacher trainings, Teacher coaching, GEFs, Mentors, CECs coaching, MOE consultations/supervisors/TOTs, Introducing Score cards for schools in the use of positive discipline, digital platform, ESL, ALP Modules, etc.
- Print and distribute the code of conduct for teachers in all schools.

The following adaptations were adopted in light of project monitoring:

- Adaptation 1 : Exclusion of OOS girls who have never been to school and those dropped out in Grade 1 or 2 or 3
  - Explore the possibility of running ABE classes for OOS girls who reached utmost Grade 3.
  - Ensure ALP students of schooling age and who have not dropped out for long time (less than 2 years) to enrol back to formal schools.



- Engage CECs to mobilise resources for an additional teacher salary/incentive.
- Expand the ALP to other villages where there is need.
- Adaptation 2: Improving the capacity of ALP teachers to teach all 4 ALP subjects with special emphasis placed on the delivery of English lessons
  - Engage additional facilitator with complementary skills
  - Ongoing coaching with emphasis on tailor made gaps to improve classroom practice
- Adaptation 3: Reach out to nomadic and pastoralist households
  - Mapping the spread of nomadic and pastoralists households and develop appropriate strategies for reaching to sub populations to ensure inclusion. Engaging CEC and host communities to intensify efforts to enrol girls from nomadic households as well those from pastoralist families.
- Adaptation 4: Promote the use of local and cultural appropriate visual aids
  - Encourage teachers to use locally available teaching and learning materials e.g pebbles, animal droppings, sticks, bottle caps etc to promote participatory learning.
  - Support teachers and students both to draw picture related to math subjects and post to class walls (provision of manila papers, drawing color pencils, postures etc)
- Adaptation 5 : Increasing student contact time/exposure
  - Address multiple grading/ Split Classes - Due to learners being at different levels [especially in NFE, and Mathematics] Proper time allocation for each level for instance if the NFE class is one hour. The Level one should have 1 hour separately and then level two will start their period after level 1 lesson instead of just one hour multigrade session.
  - The project will discuss with MoE, CECs and school administration about the possibility of having multiple shift (morning and afternoon sessions for separate groups of students) resources allowing in over enrolled schools with CECs and school administration
- Adaptation 6: Promote the use of VSLA loans and social funds to primarily focus on supporting girls education
  - Set off plan to make awareness to VSLA groups about the importance of girls' education through village event and gathering, and using VSLA as a platform for discussing issues of education equity and socio-cultural norms affecting learning and transition.
  - Create visibility posters for all villages with VSLA groups on social funding borrowing from best practice in livelihood projects.
  - Linkage to VSLA group to school CECs, Administration
  - Encourage participants to support their children education by utilizing VSLA loans.
  - Document successful group and the evidence of how the VSLA mothers used their loan, and if any portion is used for girls education.
- And to address baseline findings related to marginalization:
  - Expand ALP to enrol further 2,345 out-of-school girls within an additional 34 villages, increasing ALP coverage from 76 villages to 110 villages
  - Provision of two year Alternative Basic Education (ABE) classes for 2,029 marginalised girls and link them with existing schools to join formal education depending on learning achievement
  - Training CECs across 199 villages in identification of different type of disabilities and support to girls and boys with disabilities,

- Work with CECs to liaise with parents of displaced out of school girls and girls with disabilities, provide targeted social support and track their attendance.
- Assess girls with disabilities for placement in regular schools or referrals to special needs facilities.
- Provide specialised equipment and learning materials for 300 girls with disabilities. Assistive devices will be provided on the basis of need and may include mobility aids, hearing aids, wheelchairs, glasses and walking equipment. Teaching-learning materials will include large-print text books, pen grips/holders, and equipment for cognitive and perception development.
- Train MoE staff and teachers in 199 schools to identify and support girls with disabilities, and train at least 110 teachers and MoE staff on inclusive and special needs education. The training will include basic special education; identification and basic assessment of girls with disabilities; building inclusive classroom environments; guidance and counselling. The training will include residential training and follow up on-site sessions.
- Encourage girls and boys from pastoralist families to participate in empowerment forums to enhance their confidence and address negative stereotypes associated with their way of living.
- Lead annual social mobilisation campaigns in 70 villages to encourage pastoralists to bring their children to school and actively participate in their education.
- Provide psycho-social counselling for development of self-esteem among girls with disabilities, and treatment of anxiety and depression
- Work with CECs and teachers to address corporal punishment, particularly against over-age and displaced adolescents and those who are struggling to learn, and promote community-managed self-monitoring of community efforts in addressing corporal punishment. Encourage teachers to employ positive disciplinary measure to deter corporal punishment.
- Incorporate sessions on identification and support for girls with disabilities in stakeholder, as well as NFE and VSL groups.
- Provide support to VSLA groups to start business upon completion of the VSLA cycle through competitive selection of most viable business ideas.
- Reinforce and encourage CECs to continue supporting need based "tuition waiver".
- Increase reading time by establishing and supporting community managed reading clubs associated to GEFs/BEFs and promote the use of culturally appropriate local learning materials.

## Assumptions

The success of SOMGEP-T is predicated on a number of assumptions which have affected, and will continue to affect, the ability of project staff to carry out, monitor, evaluate, and effect change through project activities. The project's major assumptions include:

- Schools remain open during most of the year; absence of major disruptions (widespread conflict, famine, political disturbances, economic shocks)
- Most schools adhere strongly to the intervention procedures and protocols, ensuring fidelity of implementation
- Project partners adhering to implementation guidelines/protocols
- MoE efficiency
- ALP acceptance
- High retention of out of school girls
- Complementary emergency support in case of severe drought
- Timely deployment of ALP Facilitators for ALP/ teachers are available



- ALP curriculum includes skills considered as relevant in the local job market/ businesses
- Local authorities and religious leaders are supportive
- No major disruptions to government functionality post-elections, allowing for timely implementation
- Absence of major economic shocks
- Absence of major disasters and widespread conflict
- Parents are supportive of girls' participation in GEFs and BEF's

It should be noted that at least two of these assumptions articulated as part of the project MEL-F were violated during the initial phase of implementation, with significant implications for the efficacy of project interventions. Most notably:

- Schools in conflict-affected areas closed for significant periods of time during the past year, resulting in potentially attenuated learning outcomes for all students attending those schools.
- The aftermath of the drought has resulted in ongoing economic hardship for vulnerable households, particularly pastoralists who in many cases lost their main source of livelihood as a result of the drought and are now struggling at the household level with basic issues of food insecurity.

The table below links each intervention to specific intermediate outcomes and provides a comprehensive explanation of how these will in turn contribute to achieving SOMGEP-T's long-term outcomes of learning, transition, and sustainability.

*Table 1: Project design and intervention*

<b>Intervention types</b>	<b>What is the intervention?</b>	<b>What Intermediate Outcome will the intervention will contribute to and how?</b>	<b>How will the intervention contribute to achieving the learning, transition and sustainability outcomes?</b>
<b>Improving access to post-primary options</b>	Work with MoE to develop and implement ALP	Attendance, retention, and life skills development. ALP offers out of school girls and students who are unable to/ do not wish to attend formal secondary school with an alternative option, thereby encouraging them to remain in school. The program will focus in part on developing life skills that will be relevant to the job market.	By offering an alternative pathway for girls who may have otherwise dropped out, transition rates will improve. Girls will have increased exposure to higher learning, which will boost learning outcomes. ALP's particular focus on developing life skills will ensure this intervention produces sustainable outcomes, or outcomes that are relevant to the individual and community. <sup>27</sup>
	Develop girls' life skills in upper primary through ALP, including	Life skills development. Girls will learn relevant life skills that will not only	The project's learning outcomes are focused on literacy, numeracy,

<sup>27</sup> The project is working closely with the MoEs to develop the ALP model and policies related to non-formal education, thus building the foundation for the future replication of the model through government and partner-led efforts. The ALP is directly aligned with key objectives of the ESSPs to increase enrolment and provide alternative learning opportunities for marginalized groups of girls, particularly those who dropped out after early primary.

Intervention types	What is the intervention?	What Intermediate Outcome will the intervention will contribute to and how?	How will the intervention contribute to achieving the learning, transition and sustainability outcomes?
	leadership skills, financial literacy and business selection and management of income generation activities; participation in Girls' Empowerment Fora	boost their learning outcomes and attendance, but will also enable them to contribute to the local economy once they leave school.	and financial literacy. This intervention is designed to boost these specific learning outcomes, as well as increasing the likelihood of transition into ALP or secondary education. Additionally, the focus on leadership skills and other skills relevant to the job market contributes to the sustainability of SOMGEP-T.
	Develop CECs to improve retention and transition	Attendance and retention. The enhanced capacity of CECs will enable them to develop context-appropriate strategies for improving retention and transition, which will in turn have a positive effect on attendance rates.	A focus on retention and transition is expected to have a direct impact on transition rates and learning outcomes, as girls will have better access to higher education levels. The focus on the community level will ensure buy-in and contribute to the project's sustainability at the community level.
	Provide partial grants to girls from poor families	Attendance and retention. Poverty is one of the leading reasons parents are unable to send their children, and girls in particular, to school. Providing partial grants to girls from poor families will alleviate some of the financial burden impoverished families face in sending their children to school.	Increased attendance and retention is expected to improve transition rates and learning outcomes, as girls who are in school and are properly equipped are more likely to succeed. Girls from poor families who may not have otherwise had access to education will be better equipped to participate in decision-making and economic activities.
	Equip and enrol girls in 2 boarding schools	Attendance and retention for girls living in remote areas where upper grades are not available within a short distance from villages. Many families are unable to afford the fees	Increased attendance and retention is expected to improve transition rates and learning outcomes, as girls who are in school and are properly equipped are more likely to succeed.

Intervention types	What is the intervention?	What Intermediate Outcome will the intervention will contribute to and how?	How will the intervention contribute to achieving the learning, transition and sustainability outcomes?
		<p>associated with sending their children to school, including fees associated with school enrolment, textbooks, uniforms, and other supplies. By equipping and enrolling girls in boarding schools, the burden families face will be alleviated, and girls will have the equipment they need to remain in school and succeed.</p>	<p>Girls from poor families who may not have otherwise had access to education will be better equipped to participate in decision-making and economic activities.</p>
<p><b>Supportive school practices and conditions for marginalised girls</b></p>	<p>Train teachers on improved delivery of literacy and English language, supported by digital content in all 148 primary and 55 secondary schools</p>	<p>Improved quality of teaching. Qualified teachers are in low supply in all project areas. Teacher trainings will develop the skills of teachers, thereby improving their teaching quality; increased student performance and motivation is likely to have a positive effect on attendance.</p>	<p>Improved teaching quality contributes to enhance learning and transition outcomes, as children are equipped with the literacy skills in Somali and basic English skills necessary to progress to higher levels of education. Interventions focused on improving teaching quality are expected to boost transition rates and learning outcomes in a sustainable way, by equipping children with the skills they need to succeed not only in school, but outside school as well.</p>
	<p>Train teachers on improved delivery of numeracy in all 148 primary and 55 secondary schools</p>	<p>Improved quality of teaching, addressing specific gaps. Qualified teachers are in low supply in all project areas. Teacher trainings will develop the skills of teachers, thereby improving their teaching quality.</p>	<p>Poor teaching quality contributes to poor learning and transition outcomes, as children are not equipped with the basic numeracy skills, necessary to progress to higher levels of education and to develop financial literacy. Interventions focused on improving teaching quality are expected to boost transition rates and learning outcomes in a sustainable way, by</p>

Intervention types	What is the intervention?	What Intermediate Outcome will the intervention will contribute to and how?	How will the intervention contribute to achieving the learning, transition and sustainability outcomes?
			equipping children with the skills they need to succeed not only in school, but outside school as well.
	Train teachers to provide structured remedial support to students at primary and secondary level	Improved quality of teaching. Qualified teachers are in low supply in all project areas. Teacher trainings will develop the skills of teachers, thereby improving their teaching quality.	Poor teaching quality contributes to poor learning and transition outcomes, as children are not equipped with the literacy, numeracy, and English skills necessary to progress to higher levels of education. In relation to this intervention in particular, students are more likely to drop out if they do not have proper support. Interventions focused on improving teaching quality are expected to boost transition rates and learning outcomes in a sustainable way, by equipping children with the skills they need to succeed not only in school, but outside school as well.
	Train and coach teachers to deliver the ALP curriculum	Improved quality of teaching and life skills development. Qualified teachers are in low supply in all project areas. Teacher trainings will develop the skills of teachers, thereby improving their teaching quality. Additionally, the ALP curriculum offers life skills development.	Offering an alternative pathway will increase transition rates and boost learning outcomes by keeping girls in school. Ensuring the proper delivery of the ALP curriculum, which includes a focus on life skills development, will make the intervention relevant to students and the community and contribute to its sustainability.
	Construct additional classrooms in remote primary schools; build water facilities in new secondary schools; and	Attendance and retention. Lack of infrastructure is a major issue facing marginalised	Boosts to attendance and retention are expected to contribute to improvements in transition and learning

Intervention types	What is the intervention?	What Intermediate Outcome will the intervention will contribute to and how?	How will the intervention contribute to achieving the learning, transition and sustainability outcomes?
	provide solar chargers for mobile devices/tablets and sanitary pads to schools	communities in particular. Lack of proper facilities makes it difficult for students to attend and learn well in school, particularly when schools face an increase in enrolment. Additionally, girls who do not have access to sanitary pads are more likely to stay home, or drop out of school entirely. Therefore, this intervention is expected to boost attendance and retention.	outcomes. Infrastructure development will benefit not just the current cohort of students with which SOMGEP-T is engaged, but will also benefit future students.
	Incorporate life skills and financial literacy training into GEFs and BEFs	Life skills development. This intervention is focused on providing relevant life skills training through community-based forums, enhancing attendance and learning (through increased participation in class and enhanced financial literacy skills).	Financial literacy training is one of the specific learning outcomes SOMGEP-T is expecting to influence. Financial literacy and life skills training will increase the likelihood of girls succeeding in higher levels of education, and will also equip them to contribute to the local economy through income-generating activities. These skills are expected to increase the relevance of education for students and families. Life skills – specifically leadership skills – are expected to boost students’ voice and self-confidence, enhancing classroom participation among girls.
	Provide career guidance in schools	Life skills development. Providing career guidance will help develop an appropriate support system for girls and will encourage them to seek out ways in which to achieve their future career goals.	Encouraging girls to think about their futures and how to achieve their aspirations will impress on them the importance of knowledge and education. It will also give them a clear

Intervention types	What is the intervention?	What Intermediate Outcome will the intervention will contribute to and how?	How will the intervention contribute to achieving the learning, transition and sustainability outcomes?
			pathway to achieving their goals.
<b>Positive shifts on gender and social norms at community and individual girl level</b>	Engage community-level stakeholders including religious leaders, women’s groups, men and boys	Attendance and retention. Gender and social norms are a major barrier to girls’ education. Gender norms such as those that keep girls at home helping their mothers with chores negatively affect attendance and retention rates. Through engaging with community-level stakeholders, the project will contribute to community-level understanding of the importance of girls’ education.	Boosts to attendance and retention are expected to contribute to improvements in transition and learning outcomes. Shifts in gender and social norms are expected to have a long-term, sustainable impact on the communities in which SOMGEP-T will operate.
	Expand and strengthen GEFs and create BEFs to develop leadership and mentorship skills	Life skills development, attendance and retention. In addition to providing life skills development, GEFs and BEFs will be engaged in participatory tracking of graduates during project implementation, which will help them assess their own progress in increasing transition rates.	Girls who receive leadership and mentorship skills through life skills development will be better equipped to participate in class, breaking traditional norms that restrict girls’ voice; to engage in the local economy; and to contribute to their communities in the future. Additionally, the capacity of GEFs and BEFs to track attendance and retention rates will contribute to improvements in learning and transition outcomes, and will encourage community-based organizations to think about how their actions have a direct effect on important student outcomes.
	Provide adult literacy and financial literacy classes for mothers	Attendance and retention. Evidence from SOMGEP indicates that literate mothers are	Boosts to attendance and retention are expected to contribute to improvements in

Intervention types	What is the intervention?	What Intermediate Outcome will the intervention will contribute to and how?	How will the intervention contribute to achieving the learning, transition and sustainability outcomes?
		<p>supportive of their daughters spending time with their schoolwork at home, and are also more likely to appreciate the importance of girls receiving an education. Mothers who place a higher value on education are expected to understand the importance of enrolling their girls in school and encouraging them to remain in school.</p>	<p>transition and learning outcomes. Shifts in gender and social norms are expected to have a long-term, sustainable impact on the communities in which SOMGEP-T will operate.</p>
	<p>Support the financial empowerment of mothers through savings groups (VSLA), business selection, and business coaching and mentoring</p>	<p>Attendance and retention. Female heads of household are often struggling to meet the financial and opportunity costs of education, affecting girls' attendance. Mothers who participate in VSLA are able to access funds to build small businesses and support their children's education, and are also more likely to appreciate the importance of girls receiving an education. Mothers who place a higher value on education are expected to understand the importance of enrolling their girls in school and encouraging them to remain in school.</p>	<p>Boosts to attendance and retention, linked to increased financial capacity of vulnerable households, are expected to contribute to improvements in transition and learning outcomes. Shifts in gender and social norms are expected to have a long-term, sustainable impact on the communities in which SOMGEP-T will operate.</p>
<p><b>Enhanced MoEs' capacity to deliver quality and relevant formal and informal education</b></p>	<p>1) Strengthen Gender Departments' capacity to improve girls' education outcomes through trainings, development of action planning</p>	<p>Improved school governance, quality of teaching, retention, attendance, and life skills development. Enhancing the capacity of MoEs to develop plans, administer trainings, and provide incentives will contribute to all four intermediate outcomes</p>	<p>Enhancing the capacity of MoEs to take action on girls' education will have long-term effects on the communities in which SOMGEP-T operates. It will encourage positive shifts in gender and social norms, and will give MoEs actionable ways to</p>



Intervention types	What is the intervention?	What Intermediate Outcome will the intervention will contribute to and how?	How will the intervention contribute to achieving the learning, transition and sustainability outcomes?
	<p>and provision of incentives to retain the gender focal points especially in rural areas</p> <p>2) Provide support to Regional Education Officers (REOs) and District Education Officers (DEOs) to mainstream improved teaching practices and address retention/ transition</p>	<p>by sending a strong, positive message about the importance of girls' education from the government, and by giving the government clear and actionable ways to contribute to positive changes in girls' education outcomes.</p>	<p>contribute to improving learning and transition outcomes.</p>
	<p>Support quality assurance and standards (QAS) functions at all MoE levels</p>	<p>Quality of teaching, school governance, attendance, retention, and life skills development. Enhancing the ability of MoEs to monitor and evaluate their actions will enable them to understand the current educational situation and develop effective plans for addressing any gaps that exist.</p>	<p>Enhancing the capacity of MoEs to take action on girls' education will have long-term effects on the communities in which SOMGEP-T operates. It will encourage positive shifts in gender and social norms, and will give MoEs actionable ways to contribute to improving learning and transition outcomes.</p>
	<p>Provide support to Regional Education Officers (REOs) and District Education Officers (DEOs) to mainstream improved teaching practices and address retention/ transition</p>	<p>Improved quality of teaching, school governance, attendance, retention. This intervention is focused specifically on increasing the capacity of officers who have more direct oversight over the education system in their areas to address issues related to attendance and retention and mainstream improved teaching practices.</p>	<p>Enhancing the capacity of MoEs to take action on girls' education will have long-term effects on the communities in which SOMGEP-T operates. It will encourage positive shifts in gender and social norms, and will give MoEs actionable ways to contribute to improving learning and transition outcomes.</p>
	<p>Work closely with MoE on NFE for mothers and</p>	<p>Life skills development, attendance and</p>	<p>Enhancing the capacity of MoEs to take action</p>



Intervention types	What is the intervention?	What Intermediate Outcome will the intervention will contribute to and how?	How will the intervention contribute to achieving the learning, transition and sustainability outcomes?
	entrepreneurships skills for girls	retention. Encouraging and equipping MoEs to engage with mothers and girls will have a positive influence on social and gender norms, which will increase attendance and retention rates, and will contribute directly to the life skills development of girls.	on girls' education will have long-term effects on the communities in which SOMGEP-T operates. It will encourage positive shifts in gender and social norms, and will give MoEs actionable ways to contribute to improving learning and transition outcomes.
	Development of project IEC materials in conjunction with MoE for use at stakeholder advocacy and promotion events	Life skills development, attendance, and retention. IEC materials are specific knowledge products that will be shares with Parent Teacher Associations (PTA) forums, GEFs, and BEFs. These forums contribute directly to life skills development, attendance, and retention.	Enhancing the capacity of MoEs to take action on girls' education will have long-term effects on the communities in which SOMGEP-T operates. It will encourage positive shifts in gender and social norms, and will give MoEs actionable ways to contribute to improving learning and transition outcomes.

## Beneficiaries

SOMGEP-T defines marginalized girls as those who face demand-side challenges to improvement in learning and transition outcomes, including extreme poverty, pastoralism, displacement, being over age for their grade, a high degree of exposure to violence/conflict, orphan status, disability, belonging to a minority clan, and having an illiterate mother. Marginalized girls may also face limited provision of secondary education, poor infrastructure, limited access to qualified teachers, lack of remedial education for pastoralist children, and limited capacity by teachers as well as head teachers and CEC's to address poor learning outcomes, corporal punishment, absenteeism, dropout, support girls with disabilities and those suffering from anxiety and depression. The analysis to follow provides sample breakdowns by regions, grade, age, and disability, and subsequently provides a breakdown of girls' characteristics and barriers associated with educational marginalisation.

The tables in this section provide demographic information of the evaluation sample across intervention and comparison groups and across baseline and midline samples. The sample of girls and boys are shown in the tables below disaggregated by region, grade, age, and disability.

Table 2 presents the evaluation sample broken down by region and gender across the intervention and comparison groups, both in the baseline and midline assessment, and the ALP girls group in the midline assessment. Girls in ALP constitute a distinct, new population in the midline—while they are sampled from intervention areas, because they were sampled under different criteria and received different interventions than cohort girls, they are considered separately.

The baseline sample population consisted of 1,741 girls including cohort as well as non-cohort girls such as benchmark above 19 years old, transition girls between 10-19 years of age, and girls undertaking the assessment for a pilot study (benchmark only). In the midline evaluation, a total of 1,172 girls participated in the household survey and learning assessment, while 754 boys took the learning assessment alone. The second round consists of 431 cohort girls in the intervention areas and 376 in the comparison areas as in the baseline, plus 365 ALP girls who were interviewed during this round. In the midline evaluation, the survey tracked cohort girls who were in school at the baseline, and excluded out-of-school girls as well as non-cohort girls. Thus, the analysis of differences between the rounds presented in this section will be based exclusively on the sample of cohort girls that were in school at the baseline, specifically, the 807 cohort girls from both rounds. The reduction in the sample size due to the exclusion of sample points rendered inaccessible by conflict, as well as the exclusion of out-of-school girls may have affected the study's capacity to adequately capture some of the emergent, less intense trends in learning and transition and limits the ability to make inferences on subgroup patterns.

*Table 2: Evaluation sample breakdown of girls and boys (by region)*

	Baseline		Midline		
	Intervention	Comparison	Intervention	Comparison	ALP
<b>Sample breakdown (Girls)</b>					
Total Girls	431 (100%)	376 (100%)	431 (100%)	376 (100%)	365 (100%)
<b>Sample breakdown (Boys)</b>					
Total Boys	255 (100%)	255 (100%)	153 (100%)	91 (100%)	

\*Note, an asterisk in the following tables indicates changes from baseline to midline that are statistically significant at the 95% confidence level (or higher) in a regression with cluster-robust standard errors.

Table 3 shows the evaluation sample by grade and enrolment status across the intervention and comparison areas, divided by gender. There are no statistically significant differences by grade among girls between intervention and comparison schools either from the baseline and the midlines. There are significant changes in grades between baseline and midline in both intervention and comparison areas.<sup>28</sup> Girls in grade 3 (3) are only around 6 percent in both areas in the midline compared to the 30 percent of the baseline, and girls in grade 6 (7) are 17.2 percent in intervention areas and 14.9 percent in comparison areas compared to the 0.2 percent and 0.3 percent of the baseline. These results are due to the natural progression of the girls through the education system—that is, girls are expected to advance a grade after each academic year and as such significant changes in grade between baseline and midline are expected. It should be noted that while girls who were out of school in the baseline were not surveyed in the midline, there were girls who were enrolled in school in the baseline, but dropped out of school in the midline. Among the true panel of girls for the midline evaluation, 3.5 percent of them dropped out of school in intervention areas and 6.9 percent dropped out in comparison areas from the time of the baseline evaluation.

The evaluation sample of boys shows a similar pattern of change between baseline and midline assessments. There are no significant changes between intervention and comparison groups, and a significant positive correlation between rounds and grade is observed.<sup>29</sup> As per above, boys in grade 3 (3) went from making up 22.4 percent to 10.5 percent of the sample in intervention areas and from 16.5 percent

<sup>28</sup> Significance tests on baseline and midline correlation with grade provide a P-value of 0.000 for the intervention areas, and a P-value of 0.002 for the comparison areas.

<sup>29</sup> Linear regression tests result in P-values of 0.00 for both intervention and comparison areas, and a coefficient of 0.05 for both areas.

to 15.4 percent in comparison areas. Boys in grade 6 (7) compose 13.1 percent of all boys in intervention areas and 15.4 percent in comparison areas while no boy was enrolled in grade 7 during the baseline. This consistent pattern of change is likely due to the fact that at least some of the same boys assessed in the baseline were administered the learning assessment during the midline evaluation as well since the households of the cohort girls were revisited during this round.

*Table 3: Evaluation sample breakdown of girls and boys (by grade)*

	Baseline		Midline		
	Intervention	Comparison	Intervention	Comparison	ALP
<b>Sample breakdown (Girls)</b>					
OOS	0 (0%)	0 (0%)	15 (3.5%)	26 (6.9%)	365 (100%)
Unreconciled grade	2 (0.2%)	1 (0%)	5 (0%)	7 (1.1%)	0 (0%)
Primary 3	130 (30.2%)	120 (31.9%)	26 (6%)	26 (6.9%)	0 (0%)
Primary 4	102 (23.7%)	75 (19.9%)	116 (26.9%)	88 (23.4%)	0 (0%)
Primary 5	107 (24.8%)	103 (27.4%)	104 (24.1%)	72 (19.1%)	0 (0%)
Primary 6	89 (20.6%)	76 (20.2%)	87 (20.2%)	97 (25.8%)	0 (0%)
Primary 7	1 (0.2%)	1 (0.3%)	74 (17.2%)	56 (14.9%)	0 (0%)
Primary 8	0 (0%)	0 (0%)	4 (0.9%)	4 (1.1%)	0 (0%)
Total girls	431 (100%)	376 (100%)	431 (100%)	376 (100%)	365 (100%)
OOS	0 (0%)	0 (0%)	15 (3.5%)	26 (6.9%)	365 (100%)
<b>Sample breakdown (Boys)</b>					
OOS	88 (34.5%)	76 (29.8%)	26 (17%)	6 (6.6%)	
Unreconciled grade	0 (0%)	0 (0%)	12 (3.3%)	7 (5.5%)	
Primary 3	57 (22.4%)	42 (16.5%)	16 (10.5%)	14 (15.4%)	
Primary 4	31 (12.2%)	45 (17.6%)	22 (14.4%)	12 (13.2%)	
Primary 5	45 (17.6%)	46 (18%)	23 (15%)	14 (15.4%)	
Primary 6	34 (13.3%)	46 (18%)	24 (15.7%)	14 (15.4%)	
Primary 7	0 (0%)	0 (0%)	20 (13.1%)	14 (15.4%)	
Primary 8	0 (0%)	0 (0%)	2 (1.3%)	5 (5.5%)	
Total boys	255 (100%)	255 (100%)	153 (100%)	91 (100%)	
OOS	88 (34.5%)	76 (29.8%)	26 (17%)	6 (6.6%)	

Table 4 shows the population sample of girls and boys disaggregated by age groups. As per the baseline, there are no statistically significant differences in age between intervention and comparison areas. Of course, the table presents significant changes from the baseline in both intervention and comparison areas with a slight increase in age, especially regarding the sample population between 14 and 16 years old, and this trend can be again explained for the girls population evaluated in both rounds as they are expected to grow up throughout the period of the project. The same holds for the boys' sample as the same individuals have been re-evaluated after the households of the girls assessed in the baseline were re-visited.

*Table 4: Evaluation sample breakdown of girls and boys (by age)*

	Baseline		Midline		
	Intervention	Comparison	Intervention	Comparison	ALP
<b>Sample breakdown (Girls)</b>					
Aged 9-11	131 (30.4%)	112 (29.8%)	72 (16.7%)	65 (17.3%)	1 (0.3%)
Aged 12-13	144 (33.4%)	128 (34%)	147 (34.1%)	136 (36.2%)	22 (6%)
Aged 14-15	105 (24.4%)	101 (26.9%)	135 (31.3%)	107 (28.5%)	74 (20.3%)
Aged 16-17	43 (10%)	25 (6.6%)	64 (14.8%)	55 (14.6%)	100 (27.4%)
Aged 18-19	8 (1.9%)	10 (2.7%)	11 (2.6%)	11 (2.9%)	108 (29.6%)
Aged 20+	0 (0%)	0 (0%)	2 (0.5%)	2 (0.5%)	60 (16.4%)
Total girls	431 (100%)	376 (100%)	431 (100%)	376 (100%)	365 (100%)
<b>Sample breakdown (Boys)</b>					
Aged 9-11	73 (28.6%)	83 (32.5%)	27 (17.6%)	19 (20.9%)	
Aged 12-13	87 (34.1%)	94 (36.9%)	63 (41.2%)	33 (36.3%)	
Aged 14-15	59 (23.1%)	46 (18%)	37 (24.2%)	25 (27.5%)	
Aged 16-17	23 (9%)	24 (9.4%)	15 (9.8%)	9 (9.9%)	
Aged 18-19	13 (5.1%)	8 (3.1%)	8 (5.2%)	5 (5.5%)	
Aged 20+	0 (0%)	0 (0%)	3 (2%)	0 (0%)	
Total boys	255 (100%)	255 (100%)	153 (100%)	91 (100%)	

## 3. Evaluation Methodology

### 3.1. Key evaluation questions & role of the midline

Data collection for the SOMGEP-T midline evaluation began on November 6, 2018, and concluded on December 10. In the broadest possible terms, there were two primary purposes of the midline evaluation: first, to assess the project's impact after the initial exposure period of four months, by making comparisons to baseline outcomes in learning, transition, and the project's selected set of intermediate outcomes, informing

course correction of activities at this early stage; second, to serve as a baseline evaluation for a separate set of project beneficiaries – girls who are enrolled in Alternative Learning Programs (ALPs), or “ALP Girls.” In this section, we provide the methodological details necessary to understand the analysis and results that follow. Specifically, the next section provides an overview of the overall evaluation design, most of which was established prior to baseline data collection, but which informs the structure and analysis of the midline. Building on that overview, we describe the role of the midline, summarize the data collection tools employed at the midline, and list the evaluation questions that guided our analysis. More specific sampling methodology – including a discussion of the sampling adjustments made at midline and their consequences – and a description of methodological limitations to bear in mind are handled in their own discrete sections.

## Overall Evaluation Questions

The evaluation of SOMGEP-T is geared toward answering a core set of evaluation questions asked of all GEC-T projects, and an ancillary set of questions specific to SOMGEP-T and its operating environment. The core questions the evaluation seeks to address are:

- Was the GEC successfully designed and implemented? Was the GEC good Value for Money?
- What impact did the GEC Funding have on the transition of marginalised girls through education stages and their learning?
- What works to facilitate transition of marginalised girls through education stages and increase their learning?
- How sustainable were the activities funded by the GEC and was the program successful in leveraging additional interest and investment?

In addition, the evaluation seeks to answer a number of additional questions that reflect the project's specific Theory of Change, the operating environment in Somalia, and which will be used to impact future programming. These questions are:

- To what extent did the intervention result in additional gains in learning (literacy and numeracy) among the intervention group, in relation to the comparison group?
- To what extent did the intervention result in additional gains in transition to upper primary/ post-primary education among the intervention group, in relation to a benchmark sample?
- Is there a significant difference between the acquisition of literacy/ numeracy/ English skills among ASLP participants and marginalised girls enrolled in formal secondary school?
- Are the intermediate outcomes identified by the project contributing to the accelerated acquisition of literacy/ numeracy skills and improved transition rates? Are there intermediate outcomes that do not seem to be influencing outcomes at all? Do the findings support the ToC or challenge its assumptions?
- What are the key factors influencing the acquisition of literacy, numeracy and English language skills?
- What are the specific literacy/ numeracy/ English competencies that marginalised girls are lagging behind on?
- Is there a difference in the rate of acquisition of literacy / numeracy subtasks that girls are able to practice in their daily lives, vis-à-vis the acquisition rate of subtasks that are not used on a regular basis by the girls targeted by the project?
- To what extent are extremely marginalised sub-groups, such as pastoralist girls and disabled girls, able to attain basic competencies in literacy, numeracy and English? Are there other sub-groups who are demonstrating a pattern of lagging behind in performance? What are the potential reasons for this pattern?
- Is there a relationship between the acquisition of leadership skills and learning outcomes? If yes, how?
- To what extent the participation of mothers in VSL may influence the acquisition of numeracy and financial literacy skills by girls?
- What are the key factors influencing transitions to more advanced levels of education?

- To what extent are extremely marginalised sub-groups, such as pastoralist girls and disabled girls, able to transition into upper primary/ post-primary education? Are there other groups who are lagging behind in transition rates? What are the potential reasons for this pattern?
- To what extent is the acquisition of leadership skills influencing transition outcomes?
- Did the intervention contribute to a shift in traditional gender norms and power relationships at the household and community levels? If yes, what types of changes have occurred? How are these changes affecting adolescent girls and boys?
- How did the intervention affect boys' learning and retention?
- What are the key changes identified by the girls themselves in terms of their capacity to engage in non-traditional roles at the household, school and community? To what extent are those claims supported by quantitative evidence?
- Is there a difference in the learning outcomes for students targeted in areas heavily affected by drought, compared to those that were less affected? Likewise, is there any difference for transition?

The overall evaluation design, described in the next section and more extensively in the baseline evaluation report, was designed with these questions in mind.

### Overall Evaluation Design

This section provides a description of the overall research design utilized by SOMGEP-T, from baseline through the conclusion of the project. This design was decided upon jointly by Forcier and CARE, in consultation with the FM, immediately prior to the baseline round of data collection. It is largely in line with the project's Monitoring, Evaluation and Learning Framework (MELF), while incorporating changes approved by the FM in August 2018.

The overall evaluation employs a quasi-experimental difference-in-differences design for assessing the project's impact over time. The evaluation incorporates an explicit comparison group of schools that are not receiving SOMGEP-T interventions or similar interventions from other development actors; school-age girls sampled from the communities surrounding these non-intervention schools constitute the comparison group for assessment of learning and transition outcomes, as well as intermediate outcomes related to attendance, life skills, teaching quality and school governance. Data collection in both comparison and intervention communities occurred at the baseline and continued through this midline evaluation wave (though with an adjusted sample of schools or communities), a third evaluation round (second midline), and the endline. This design will allow inferences to be drawn regarding changes in girls' assessment scores and other key outcomes over time, by comparing the relative change among girls in intervention schools to the relative change among girls in control schools.

The difference-in-differences approach accounts for two important sources of bias when drawing inferences regarding project impact: first, by comparing the same respondents pre-intervention and post-intervention, the design accounts for underlying cross-sectional differences between intervention and comparison groups. For instance, if project schools already perform at a higher level than non-intervention schools at the time of the baseline or in the absence of any intervention (i.e. under the counterfactual situation where no projects were implemented in any school), this fact would bias straightforward cross-sectional comparisons of outcomes at the endline. By adjusting explicitly for gaps that existed prior to project implementation, the difference-in-differences approach ensures that pre-existing differences do not bias the conclusions drawn about project impact. Second, the design controls for systematic changes in outcomes over time that are not attributable to the project itself. For instance, in the context of a widespread shock that reduced learning outcomes across most of the region, a simple longitudinal comparison of learning outcomes in project schools over time would suggest that the project reduced student performance. By incorporating an explicit

comparison group, the design is able to control for systematic shocks that affect both intervention and non-intervention schools.<sup>30</sup>

Inferences drawn via difference-in-differences rely on two key assumptions. The first assumption is that, under the counterfactual condition – i.e. in the absence of intervention – the change in outcomes over time in intervention and comparison schools will be similar. This assumption is often referred to as the “parallel trends assumption,” because it assumes that – in the absence of intervention – trends in outcomes in the intervention group would evolve in a fashion parallel to those in the comparison group.<sup>31</sup> The second assumption is that comparison group respondents are not exposed to the intervention or to similar interventions affecting learning and transition. Informally, this assumption is often stated as the “no spillover” assumption.

The first assumption is, in general, untestable under a difference-in-differences framework. In the absence of randomization, the parallel trends assumption may or may not hold, though the selection process employed by SOMGEP-T leaves no reason to *expect* divergent trends under the counterfactual. The second assumption is testable *ex post*, by directly measuring whether comparison group respondents received any of the project’s interventions. If both assumptions are satisfied, inferences drawn using difference-in-differences are unbiased in expectation.

Beyond the overall evaluation setup, it is important to note that SOMGEP-T employed a joint sampling approach when selecting girls for inclusion in the cohort at baseline. In a joint sampling approach, the girls selected at the baseline constitute a cohort to be tracked over time for the purposes of assessing *both* learning and transition outcomes.

During the baseline, the evaluation team sampled girls at their households, rather than at the project’s schools. This sampling approach avoided the significant bias that would occur if girls were selected at schools, as only in-school girls would have been included in the latter case.<sup>32</sup> Because girls were selected at their households, the sample included a considerable number of out-of-school girls; moreover, the approach allowed the evaluation team to collect data on the household characteristics of girls in the cohort.

The design described here applies to the overall evaluation methodology for SOMGEP-T and has not been altered from the initial decisions made prior to the baseline. With that said, the methodology specific to the midline has been adjusted in places, including in the population of cohort girls studied. In the next section, we discuss midline-specific methodological adjustments.

## Role of the Midline Evaluation

This midline evaluation constitutes the first effort at measuring the impact of SOMGEP-T at its initial implementation phase. As noted above, the midline evaluation serves two primary purposes. The first purpose is to assess the impact of SOMGEP-T after four months of exposure on learning, transition, and intermediate outcomes, by comparing girls and schools in intervention and control communities. In general,

---

<sup>30</sup> Put in different terms, the difference-in-differences design accounts for bias from heterogeneity between intervention and control groups (pre-existing differences between the two) and from temporal shocks that impact both intervention and control groups equally. Bias is still possible, however, as a result of temporal shocks that are specific to either the intervention or control group or disproportionately affect only one of these groups. Such bias is a weakness of all designs including, to some degree, randomized controlled trials (RCTs) or experiments. Bias of this form can be controlled using standard multivariate regression techniques, if data on the nature and extent of the shock is available.

<sup>31</sup> In line with the discussion above, this assumption means that there are no temporal shocks that apply disproportionately to only the intervention or control group.

<sup>32</sup> In fact, the bias is even greater than simply the exclusion of out-of-school (OOS) girls. Sampling at schools also ensures that girls who *attend* school most often have a higher probability of selection, resulting in a sample that is heavily biased toward both enrolled girls and the girls who attend school most frequently.



this analysis focuses on relative change over time, comparing the evolution of, for instance, learning outcomes in intervention and comparison communities from baseline to midline. In short, the analysis employs a classic difference-in-differences framework, as the overall evaluation design intended.

The second purpose is to provide a baseline evaluation of formerly out of school girls who are now enrolled in Alternative Learning Programs (ALPs). Girls who are enrolled in ALPs are fundamentally different from the cohort girls selected at baseline. ALP girls have enrolled in alternative education and tend to be older than the typical cohort girl. This evaluation serves as a baseline for these girls, with the goal of describing them and understanding the unique challenges they face. In addition, the data collected will serve as a baseline against which their progress can be tracked in future evaluation waves.

## Data Collection Tools

The midline is a mixed-method evaluation, employing both quantitative and qualitative data collection tools. As with all impact evaluations, the extent to which a particular indicator is targeted with qualitative versus quantitative data varies depending on its nature. For instance, learning and transition outcomes are predominantly addressed through the use of quantitative data, given that qualitative assessments of learning performance are difficult to use. Even here, though, the results are contextualized extensively with qualitative data, particularly with regard to the barriers and challenges that prevent learning or shape transition outcomes. In other areas, particularly some indicators of sustainability, the evaluation makes extensive use of qualitative data.

The quantitative tools employed at the midline include:

- Household survey and learning assessments
  - Girl module – life skills, self-esteem, teaching quality, disability status, etc. including the Youth Leadership Index (YLI)
  - Head of household module – household characteristics
  - Caregiver Module – girls' enrolment status, teaching quality, school management, disability status, etc.
  - Learning assessments – numeracy (SeGMA) and Somali literacy (SeGRA)<sup>33</sup>
- Head teacher survey
  - Head teacher module – school environment, infrastructure, teacher quality, etc.
  - School records module – measures attendance for cohort girls using school records
- Classroom observations
- Attendance spot checks or classroom headcounts

From the perspective of the quantitative tools, the most significant change from baseline to midline is the removal of the English assessment from the SEGRA and the exclusion of a dedicated teacher survey during this round of data collection. The baseline evaluation included a self-administered survey of 436 teachers, which collected data on teachers' attitudes toward girls' education and their teaching practices, among other things. This information was not included in the midline, though it may be utilized in future evaluation waves. Importantly, some outcomes – especially those focused on teaching quality and CEC financial support of schools – are adversely affected by the lack of data from teachers.

In terms of qualitative data collection, the midline evaluation included interviews and focus groups with a diverse array of beneficiaries and stakeholders. The specific qualitative tools employed were:

- FGD with Community Education Committee (CEC) members
- KII with Ministry of Education (MOE) officials
- FGD with teachers
- FGD with female community members
- KII with girls with disabilities

---

<sup>33</sup> The English component of the SEGRA was not included in this evaluation round, because students had not yet been exposed to project interventions linked to English literacy at the time of data collection.



- Participatory risk mapping with girls who are members of the Girls Empowerment Forum
- Participatory story-telling exercises (vignettes) with girls who are members of the Girls Empowerment Forum

Compared to the quantitative tools, more extensive changes were made to the qualitative tools. Specifically, three new data collection instruments – KIIs with girls with disabilities, participatory risk mapping, and participatory story-telling exercises – were incorporated into the midline. The goal of these new tools was to increase the visibility of girls’ own views and voices in the final report. While the baseline included FGDs with male and female students, they did not specifically target girls with disabilities, and they covered a wide range of topics. The midline’s inclusion of a dedicated tool allowed girls with disabilities to be heard more clearly, with questions designed specifically to elicit insights into the unique challenges they face. The participatory exercises, meanwhile, facilitated data collection on targeted areas of interest: safety, child protection, and comfort in their communities in the case of risk mapping, and a broad array of barriers to educational attainment, in the case of participatory story-telling.

Prior to analysis, all qualitative data was transcribed and translated from Somali into English. Transcription, translation, and quality control checks of translation were all conducted by Forcier’s core national staff. Qualitative data was analysed systematically, employing an exploratory coding scheme. After multiple reviews of the qualitative data, themes were allowed to emerge based on the respondents’ own stated concerns, sentiments, explanations and insights. These emergent themes were coded broadly, independent of the quantitative findings, with qualitative analysis initially conducted by a dedicated analyst. Coding was gradually refined until potentially generalizable insights could be extracted regarding hypothesized relationships of interest to the evaluation. Emergent themes were also identified that can speak to future program/intervention priorities, potentially unforeseen causes or consequences of interest, as well as previously unidentified gaps in programming.

As a second layer of analysis, the qualitative data was examined for narrative evidence that could shed light on the lived experiences behind the quantitative findings, including prevalent social and gender norms (which may not be expressed in quantitative surveys as a result of social desirability bias, but surface in qualitative narratives). Counter-narratives or minority narratives (that potentially contradict or qualify quantitative findings) were also given voice, to the extent possible. Finally, qualitative data was often queried selectively to make sense of or contextualize specific quantitative findings that were either outliers or counter-intuitive in some way. Most importantly, the initial qualitative analysis was explicitly separated from the quantitative analysis, so that qualitative findings could emerge without reference to specific quantitative results. Only after these initial findings were described did the qualitative analyst review the quantitative results; at that stage, they moved to contextualization and explanation of the quantitative findings, while still preserving the broad qualitative themes from their initial analysis.

Additionally, qualitative data was used to inform the analysis on specific qualitative indicators included in the project logframe. Outlined in the table below are the intermediate outcomes that have qualitative indicators, the indicators themselves, the midline targets for the qualitative indicators, and the qualitative tools used to inform the analysis on each indicator. It should be noted that, given the “light” qualitative approach taken in the midline, FGDs with girl and boy students were not included in the sampling approach as they were in the baseline. Therefore, for the life skills outcome, girls’ experiences were directly captured primarily in the interviews with girls with disabilities, or were inferred from information shared by teachers and mothers. Girls were also directly interviewed in the participatory exercises – the risk mapping and vignettes – but these exercises were not designed to elicit information on the life skills indicator, which is why these tools are not

Intermediate Outcome	Indicator	Midline Target	Qualitative Tool(s) Informing Analysis
----------------------	-----------	----------------	--

<b>Attendance</b>	Mothers' support to adolescent girls' attendance.	Mothers express support to adolescent daughters' attendance to ASLP and formal school.	<ul style="list-style-type: none"> <li>• FGDs with mothers</li> <li>• FGDs with CEC members</li> <li>• IDIs with girls with disabilities</li> <li>• KIIs with REOs</li> </ul>
<b>School governance</b>	CECs' perceptions of the importance of the retention of marginalised sub-groups, such as pastoralists.	CECs include retention of marginalised groups in school improvement plans.	<ul style="list-style-type: none"> <li>• FGDs with CEC members</li> <li>• FGDs with teachers</li> <li>• IDIs with girls with disabilities</li> </ul>
<b>Teaching quality</b>	Shifts in teachers' awareness of quality of education.	Teachers are aware of the need to support students who are lagging behind in acquiring literacy and numeracy skills; identify sub-groups who are struggling and potential strategies for inclusion.	<ul style="list-style-type: none"> <li>• FGDs with teachers</li> <li>• IDIs with girls with disabilities</li> </ul>
<b>Life skills</b>	Girls feel comfortable expressing themselves at school, in the community and at home.	Girls describe examples of engagement with others at school and household to express their needs and aspirations.	<ul style="list-style-type: none"> <li>• FGDs with mothers</li> <li>• FGDs with teachers</li> <li>• IDIs with girls</li> </ul>

### Midline Evaluation Purpose

As noted above, the SOMGEP-T midline evaluation serves two distinct purposes, for two distinct subgroups of beneficiaries. With regard to cohort girls broadly and the communities in which they live, the midline set out to evaluate the project's impact over the last year in terms of the GEC-T core outcomes and intermediate outcomes. Specifically, the evaluation assesses impact on the intervention group from baseline to midline in terms of:

- Learning scores (numeracy and Somali literacy)
- Transition, including in-school progression, enrolment in alternative education or vocational training, or age-appropriate employment
- Attendance (IO 1)
- School management and governance (IO 2)
- Teaching quality (IO 3)
- Life skills and self-esteem (IO 4)

In some cases, the analysis moves beyond aggregate project impact and also considers heterogeneous impact or intervention effects by, for instance, assessing project impact in specific subsamples of communities or respondents (e.g., project impact on learning outcomes among grade 4 students).

The evaluation also sought to assess the project's likelihood of being sustained over time, following the conclusion of programming. The project's MEL Framework has identified four indicators of sustainability; while the midline did not always seek to formally assess impact on these indicators (i.e. the differential

evolution of the indicator in intervention versus comparison communities), they were assessed broadly to score the project's current level of sustainability:

- CEC financial support for schools
- Adherence to implementation standards for ALP, numeracy and remedial classes
- MOE engagement with and support of girls' education initiatives
- Inclusion of ALP in national non-formal education frameworks

With regard to ALP girls, for whom this evaluation serves as a baseline assessment, three core evaluation questions defined the analysis:

- How do ALP girls perform in terms of learning outcomes and life skills at the baseline?
- What barriers or characteristics influence learning outcomes specifically among ALP girls?
- How do ALP girls and their households differ from girls who are enrolled in school and those who dropped out following the baseline but are not enrolled in alternative education?

The first ALP evaluation question is focused on establishing baseline values against which future progress can be measured. The second evaluation question will allow CARE to identify the type of ALP girls that are at particular risk or who underperform relative to their peers, and may facilitate adjustments to programming in response. The third evaluation question considers ALP girls relative to their non-ALP counterparts, providing insight into *why* some girls enrol in ALP and others do not, with the goal of identifying patterns that could be used to adjust project implementation (by, for instance, targeting for ALP recruitment OOS girls who are particularly likely to enrol, given their characteristics).

## 3.2. Sampling Methodology

### Sampling and Re-contact of Cohort Girls

The SOMGEP-T baseline evaluation selected a cohort of girls to be tracked longitudinally using a joint sampling approach, as outlined above and – in more detail – in the baseline evaluation report. Two important changes to the sample design were made at the midline. The first concerns the sample of schools which will be visited; the second concerns the set of girls in the baseline cohort who will be re-contacted.

At the midline, the evaluation team visited a truncated set of the same intervention and comparison schools from the baseline. At the baseline, 76 schools were visited in total, with the sample evenly split between intervention and comparison schools, which were matched using Coarsened Exact Matching (CEM) to create a sample balanced on pre-intervention observable characteristics, to the extent possible. At the midline, the evaluation team re-visited 63 of the same schools.

Notably, five schools were excluded from the midline sample because they were outliers – in terms of learning outcomes, especially – at the baseline. The five schools had particularly skilled teachers with regard to English-language instruction, and were typically removed from consideration during the analysis of the baseline data (or different sets of results – some including them and some excluding them – were reported). Note that the excluded schools can be reincorporated into the sample in future evaluation waves.

The second sampling adjustment concerns the set of cohort girls who were re-contacted at the midline. At included schools, the evaluation team re-contacted a subset of the girls' learning and transition cohort (i.e. the baseline sample of girls). Specifically, the midline sought to re-contact only girls who were enrolled in school at the time of the baseline. Girls who were out-of-school at the baseline – whether enrolled in alternative education or employed or not otherwise engaged – were not re-contacted in this round of data collection. Importantly, all members of the learning and transition cohort will be re-contacted at the third evaluation round and endline for the purposes of assessing aggregate transition and learning outcomes over the life of the project.

Naturally, these sampling decisions have important consequences for the types of analysis that can be conducted in this round. First, all analysis that makes comparisons between baseline and endline – including comparisons of learning and intermediate outcomes – uses the subset of respondents or observations in the baseline that are comparable to the midline. This approach is simplest in the case of school- or classroom-level analysis, such as the analysis of the head teacher survey, headcounts, or classroom observations. In these cases, we subset the baseline sample to include only those schools that also appear in the midline sample. The population being studied has, therefore, changed: when we assess a change in attendance rates via classroom headcounts, we can only say that the change in attendance rates from baseline to this second evaluation round applies to the subsample of schools that was included in both waves.

In the context of girl- or household-level analysis, the picture is complicated by the fact that some girls targeted for re-contact and inclusion in the midline sample were not successfully located and had to be replaced. Effectively, there are two different samples of girls that are arguably comparable from baseline to midline. “True panel girls” are – as the name implies – those who constitute a true panel dataset. These are girls who appeared in the baseline, were enrolled in school at that time, whose schools are also included in the midline, and who were successfully re-contacted. Comparisons using this panel are the most rigorous in terms of avoiding bias due to sampling variance from baseline to midline, because the set of girls analysed has not changed at all. “Sample girls” share many characteristics with “true panel girls”: specifically, they appeared in the baseline, were enrolled in school at that time, and their schools are also included in the midline. Where they diverge is in their re-contact status – “sample girls” include those girls who were not successfully re-contacted and had to be replaced at midline by a girl from the same school. As with the school-level analysis described above, the population of girls being analysed has changed from baseline. In every case, the population being studied is the set of girls who were enrolled at baseline and whose schools appear in both the baseline and midline samples. In the case of true panel girls, the sample includes only those who were successfully re-contacted; in the case of sample girls, the sample includes true panel girls as well as those girls who were not successfully re-contacted and were replaced instead.

The second, and arguably more important, consequence of changes to the sampling methodology concerns the measurement of transition rates. Re-contacting only in-school girls from the baseline makes it impossible to accurately measure transition rates from baseline to midline in the same manner utilized by the baseline evaluation. Among in-school girls at the baseline, the evaluation can track only *dropout rates* (i.e. the share of girls who successfully stayed in school or dropped out) and transition into non-school alternatives. But the project will not be able to assess either re-enrolment rates among OOS girls from the baseline, or the extent to which OOS girls from the baseline entered ALPs, vocational training, or employment. While re-enrolment is relatively rarer than dropping out in the first place, focusing exclusively on successful transition of in-school girls to their next grade level will overestimate transition rates, as the sampled girls are relatively more likely to successfully transition than the OOS girls included at the baseline.

To avoid this fundamental problem of incomparability, the midline evaluation will report successful transition rates among the midline sample of girls who were enrolled in school at the baseline. The comparison group from the baseline will be girls who were reportedly enrolled in school in the year *prior* to the baseline. For the sake of clarity, let the baseline be time  $t$ . During the baseline, transition outcomes were defined by changes in a girl's enrolment status from time  $t-1$  to time  $t$ . As in all GEC-T midline evaluations, transition at the midline is defined as changes in enrolment status between time  $t$  to time  $t+1$ . The core comparison in this evaluation will be midline transition (from time  $t$  to time  $t+1$ ) among those girls who were enrolled at time  $t$ , compared to baseline transition (from time  $t-1$  to time  $t$ ) among those girls who were enrolled at time  $t-1$ . Importantly, these two samples will not overlap entirely, as some girls who were enrolled at time  $t-1$  were not enrolled at time  $t$  and will be included in the calculation of baseline transition rates but not midline transition rates. However, because the set of schools will remain fixed between baseline and midline, this comparison should provide a valid comparison of broader community dropout/transition rates among in-school girls between baseline and midline.

Beyond the issue of sample selection, the cohort girl sample is influenced by the process used to locate or re-contact girls at the midline, and the process for replacing those girls who could not be re-contacted. In Section 3.4, below, we provide an in-depth analysis of re-contact or attrition rates, including the predictors of attrition, and compare replaced to replacement girls. Here, we simply describe the process by which girls were re-contacted or replaced.

Field teams were provided a list of cohort girls to be re-contacted in each of their assigned schools. Upon arrival at the school in question, team leaders met with the school's head teacher and reviewed the list of girls to be interviewed. Most girls were located successfully at this stage, either because the girl was still enrolled and present in school, or because the head teacher was familiar with the girl and knew where she could be located. When a girl could not be located at the school or via the head teacher, a specific protocol was followed for locating her, which included the following steps (though not necessarily completed in this order):

- Call the primary phone number provided by the girl's household at baseline, with a minimum of two calls made, separated by at least six hours
- Call the secondary phone number provided by the girl's household at baseline (if available), with a minimum of two calls made, separated by at least six hours
- Visit the location of the girl's household, with a minimum of two visits, separated by at least six hours
- Contact the household's nearest neighbours, if the household consented to such contact at the baseline
- Visit the local mosque or other local gathering place and ask community members about the girl and her family
- Ask girls at the school, in the same age and grade cohort as the target girl, if she still lives in the community and how she can be reached

If a girl from the baseline cohort could not be located after these steps were completed, a replacement girl was selected for the midline. Replacement girls are all enrolled in school at the midline. Where possible, replacement girls were selected randomly from a list of girls in the target girl's same grade (i.e. if she was enrolled in grade 4 at baseline, possible replacements would be limited to girls in grade 4 (5)), who are also in the project's targeted age range (11-21 years) and are not already part of the cohort sample. In cases where no girls in the same grade are available as replacements, the grade cohort restriction was widened until a suitable replacement girl was found.<sup>34</sup> Once field teams constructed a list of eligible replacement girls for each target girl who could not be located, they selected a replacement from each list using a random number generator. In total, the successful re-contact rate for cohort girls targeted at the midline was 80.3 percent, a point we cover in more detail in Section 3.4.

Replacement girls were selected both for the analysis in this midline report and to top-up the sample going forward, i.e. to provide a refreshed panel for analysis of midline to endline changes in the eventual endline evaluation. To be clear, replacement girls are often included in the analysis reported here: for instance, when analysing student perceptions of teaching quality from baseline to midline, we include both replacement (midline) and replaced (baseline) girls, alongside those girls who were successfully re-contacted and appear in both rounds. Our primary results regarding learning and transition outcomes, however, include only the "true panel" -- girls who were successfully re-contacted and appear in both rounds.<sup>35</sup> We report additional results that include replacement and replaced girls, but we clearly note where the results are based on the true panel versus the larger sample that includes replacement girls.

Replacement girls also serve to refresh or top-up the samples going forward. Both the learning and transition samples were refreshed at the midline, because a joint sampling approach was employed and the

---

<sup>34</sup> For instance, if a girl in grade 4 at the baseline could not be located, the team would first seek to replace her with a grade 5 girl, then expand the set of acceptable grades for replacement to grades 4 and 6, then expand the set of acceptable grades to grades 3 and 7, until an eligible replacement girl could be located.

<sup>35</sup> This includes the results provided in the Outcomes Spreadsheet and the project logframe.



learning and transition samples in the overall evaluation are comprised of the same set of cohort girls.<sup>36</sup> In total, 159 replacement girls were selected; out of a target sample of 824 girls, the topped-up samples comprise 807 girls, providing a mostly undiminished base of girls to re-contact and analyse for midline to endline changes. The refreshed panel of respondents will allow rigorous analysis, with a large panel sample, of learning and transition outcomes from midline to endline.

### **ALP Girls Sampling**

Alternative Learning Programs (ALPs) have not been implemented in every community where SOMGEP-T has been started. Specifically, ALPs have not been established in areas where secondary schools are available within the community or within a short distance from it. In addition, ALP girls are more difficult to locate within communities, as they represent a small subset of the population. As a result, the sampling approach for ALP girls differs substantially from that of cohort girls.

The sample was drawn from a list of girls enrolled in ALPs organized by SOMGEP-T. The sample selection process was not fully random; it included all available ALP-enrolled girls in ALP centres whose locations overlapped with the formal school sample, plus a randomly selected sample of additional ALP sites. All ALP centres that overlapped with the formal school sample were selected for logistical and budget reasons, because field teams were already traveling to those locations. The randomly selected sample of additional ALP sites was necessary to ensure a sufficiently large sample of ALP girls for analysis. In addition, some ALP centres were removed from the sample frame because they have too few participating girls or the logistical burden of reaching the site was too great.

CARE's Monitoring and Evaluation team drew the sample of ALP girls from a set of 29 locations where ALPs are active. Just over half (16) of the locations overlap with intervention schools that were visited in the baseline and midline, such that both cohort girls and ALP girls were interviewed and assessed in the same community. The remaining ALP sample locations were not included in the baseline sample of cohort schools. In total, a target sample of 354 ALP girls was identified across the 29 ALP sites. At ALP sites, all available ALP girls were interviewed (in other words, a census was conducted and girls were not randomly selected) to ensure a sufficiently large sample for analysis. This sample constitutes the ALP cohort, which will be longitudinally tracked in future evaluation waves.

### **Sampling Decision Points**

The previous two sections outline two major changes to the *structure* of the sample at this midline round. Specifically, a subset of girls – those who were enrolled at the baseline – were selected for re-contact, and girls who were OOS at the baseline will be targeted for re-contact during a second midline round. In addition, an entirely new sample of girls was added to the evaluation, consisting of girls participating in Alternative Learning Programmes (ALPs). Note that these girls were not interviewed at the baseline and are not included in the main learning or transition cohorts; they constitute a separate cohort of girls to be tracked through the endline.<sup>37</sup>

These decisions were made for a number of methodological and budgetary reasons. When the evaluation was initially designed, it included a single midline evaluation round. CARE's preference was for a second midline evaluation point for methodological reasons, but budget limitations appeared to preclude this option.

---

<sup>36</sup> We recognize that this seems to contradict our earlier discussion of the transition sample analysed in this report. In the overall evaluation, the learning and transition samples are coterminous; for the purposes of reporting transition rates at the midline, an alternative sample was constructed, owing to the unusual sampling approach taken, but the overall evaluation plan regarding the transition sample is unchanged.

<sup>37</sup> We do not discuss in detail the decision to remove schools from the sample. Individual schools were removed for security or accessibility reasons, because they were outliers at the baseline, and for other assorted reasons. These decisions potentially impact the comparability of the intervention and comparison groups, but do not fundamentally alter the sampling strategy. Our focus in this section is on decisions made that alter the structure of the sample in a deeper way.

By reducing the set of children that would be interviewed at each school, the budget limitations that prompted a single midline evaluation round were loosened sufficiently that a second midline evaluation point could be added. However, this expansion to two midline evaluation points was contingent on sampling only in-school girls (i.e. girls who were enrolled at the baseline) during the first midline – to reduce the data collection burden – and OOS girls (those who were OOS at baseline) during the second midline.

At the same time, the decision to prioritize in-school girls at midline point #1 and out-of-school girls at midline point #2 was driven by the fact that the impact on out-of-school girls from one year of implementation is likely to be fairly limited. That is, the project's impact was expected to be most immediately felt by in-school girls, who are directly targeted through schools, where project rollout can be done quite efficiently – for instance, beneficiaries can be identified more easily, school officials are already actively engaged from the first round of GEC, and the range of interventions that can quickly be rolled out is larger, among other advantages. In contrast, out-of-school girls are more difficult to target, being a diffuse group; the interventions take more time to begin implementation and ramp up to full capacity; and encouraging re-enrolment or enrolment in alternative opportunities is a more difficult outcome to influence than encouraging continued enrolment among girls already enrolled. For these reasons, it was decided that the first midline evaluation would focus on girls already enrolled in school, while the second midline evaluation would focus on out-of-school girls after allowing additional time for interventions to impact this subpopulation.

Elsewhere in this section, we have discussed the analytical consequences of these sampling decisions, whose primary impacts are on the precise nature of the transition outcome and on the sample size at midline. Somewhat less importantly, the decisions occasionally limit our ability to answer very specific questions, but otherwise do not impact our ability to analyse the intermediate or core outcomes. On a broader level, the decision to split the midline into two rounds will potentially make aggregate analysis at endline more complicated, because different girls were surveyed during different midline rounds. On the other hand, for analysis that targets girls at particular starting points (i.e. in-school versus OOS girls), the split midline will have no effect, and has the methodological and intervention-timing benefits outlined above.

The decision to sample ALP girls, in contrast, was fairly straightforward. At the time of the baseline, ALPs had not formed, at least in most communities. This midline sought to establish a cohort of girls enrolled in ALPs, for the purpose of tracking their learning and transition outcomes over time. Because ALPs had not been formed at the baseline, it would not have been possible to directly sample – via a beneficiary list, for instance – ALP girls at the baseline, whereas this was possible at the midline. As noted above, the sample of ALP girls is considered an entirely separate cohort to be tracked, and is not included in any comparisons between baseline and midline, or between intervention and comparison areas (because ALPs are present exclusively in intervention areas).

### **Overall Quantitative Sample**

Given the changes in sampling approach at midline, the overall set of quantitative observations – in every population group – has declined since baseline. The table below describes the sample targeted for inclusion at midline, and the actual achieved sample, for each quantitative data collection tool used. Two discrepancies between the targeted and achieved sample stand out. First, no specific target for classroom headcounts was established, as the sampling plan called for – in line with the baseline approach – conducting a headcount in every available classroom at each sampled school. We expected to complete between 450 and 500 headcounts, assuming that most schools have 7 to 8 classrooms; in total, headcounts were completed in 455 classrooms. Second, the number of completed household surveys and learning assessments was lower than the target. The resulting gap of 21 interviews stemmed from both fieldwork (in a few schools, no replacement girls were available, so the school-specific sample size target was missed) and from cleaning (some girls did not meet the sample's inclusion criteria of being in-school at baseline, for instance, and were removed from the sample prior to analysis).



Table 5: Targeted and achieved samples, by tool, in midline

Data Collection Tool	Sampled Population	Target Sample	Achieved Sample
<b>Cohort/Primary Schools</b>			
Household Survey	Cohort girl households	828	807
Learning Assessment	Cohort girls	828	807
Head Teacher Survey	Head teachers	63 (1 per school)	63
Classroom Observation	Classrooms	63 (1 per school)	62
Headcounts	Classrooms	450-504 (every classroom in each visited school)	455
<b>Alternative Learning Programs</b>			
Household Survey	ALP girl households	354	365
Learning Assessment	ALP girls	354	365

To make the analytical sample sizes clear for primary outcomes, the following table summarizes total sample (with replacements) and panel sample sizes for both learning and transition. It should be noted that all learning outcomes were analysed with the panel sample, by default. All results were checked for robustness against the sample with replacements. For transition, only the panel sample could be used because replacement girls have a transition rate of 100 percent by definition.

Table 6 - Learning and transition samples

Outcome	Total (with replacements)	Panel
Learning	694	564
Transition	N/A	648

*Project note:*

The “learning sample”, as referred in Table 7 above, is formed by students in grades 4, 5, 6 and 7. The analysis also includes a sample of 52 grade 3 students (demoted or repeaters) and 41 girls who have dropped out of school since the baseline; additionally, 20 students were not included in the analysis due to being in grade 8 (8) or the grade information provided could not be reconciled with records (12). Therefore, the overall number of girls participating in the learning assessment was 807.

The table below compares the baseline and midline samples in terms of specific data collection tools and targeted population groups. The table illustrates the decline in sample size for each tool from baseline to midline that results from reducing the set of schools visited. Column 2 reports the full baseline sample size for each tool; column 3 reports the “comparable” baseline sample, which has been reduced to include only those schools that were visited at the midline. The difference in sample size between columns 2 and 3 represents the decline in usable baseline sample size from restricting the set of sampled schools to match that of the midline.

Table 7: Baseline and midline samples, by tool

Data Collection Tool	Sampled Population	Full Baseline Sample	Comparable Baseline Sample	Midline Sample
Head Teacher Survey	Head teachers	78	63	63
Classroom Observation	Classrooms	152	124 (2 per school)	62 (1 per school)
Headcounts	Classrooms	506	421	455
Teacher Survey	Teachers	436	N/A	0

The change in sample size for the household survey and learning assessments conducted with cohort girls is more complex, given that it is affected by the number of girls who were enrolled at the baseline and by the number of girls who were located or replaced at the midline. The total cohort girl sample size at baseline was 1,741 girls; limiting the sample to comparable – i.e. those visited at the midline also – reduced the set of cohort girls to 1,444. Further reducing the sample to those girls who were enrolled at school at the time of the baseline produces a baseline sample size of 828 girls; once adjusted for girls who could not be located or replaced, and for those girls who did not fit the inclusion criteria produces a sample of 807 girls in both baseline and midline.

### Sampling for Qualitative Interviews

Qualitative research at the midline focuses on a diverse group of respondents, including ALP girls, CEC members, female community members, MOE officials, and girls with disabilities. The process for selecting participants in the varied qualitative interviews was designed to meet two disparate goals: first, ensure representation of a wide range of viewpoints by, for instance, targeting a variety of geographic locations and using random assignment of schools to participation in qualitative interviews where appropriate; second, ensure the quality of data by selecting schools and participants that were of particular relevance to the research goals (e.g., selecting schools from active Girls Empowerment Forums (GEFs) for participatory exercises, as GEFs are arguably the most important vector through which SOMGEP-T hopes to impact girls' life goals and self-confidence).

For logistical reasons, the following qualitative tools were grouped, such that schools/communities were selected for participation in a *cluster* of qualitative interviews:

- FGD with CEC members
- FGD with teachers
- FGD with female community members

In other words, in each of the ten communities selected for this set of qualitative interviews, field teams completed one of each of the FGDs listed above.<sup>38</sup> The ten communities were selected randomly from among the sample of intervention schools included in this midline evaluation (excluding communities where only ALPs are being targeted for evaluation, as ALPs are not managed by CECs).

Beyond the FGDs listed above, the evaluation includes KIIs with MOE officials and girls with disabilities (GWDs). Participants in these interviews were selected purposively. In the case of MOE officials, CARE provided the names and contact information for MOE officials in each region, with a total of six officials being interviewed. In the case of GWDs, CARE selected the initial set of participants from the set of cohort girls interviewed at the baseline. Specifically, CARE's Monitoring & Evaluation team selected girls who, at the

<sup>38</sup> For an overview of the data collection tools and the sample size of each type of qualitative interview, see Section 3.1.

baseline, were reported to have visual, mobility, or mental health (anxiety/depression) impairments.<sup>39</sup> Ten such girls were selected for interviewing.<sup>40</sup>

Finally, the baseline evaluation included two participatory exercises with groups of girls: risk-mapping of their communities and schools, and story-telling or vignette exercises (more details are available in Table 3). In total, eight interviews are planned – four risk-mapping and four vignette exercises. Locations were selected for the completion of the participatory exercises based on the existence of an active GEF, with assignment of interviews stratified by geographic region in proportion to the number of sampled schools in each region. Schools were chosen only if at least one girl from the baseline cohort had heard of activities in their school led by the GEF. In practice, in all but one of the eight schools, at least 20 percent of girls from the baseline cohort had heard of such activities led by GEFs. This level of activity represents relatively high GEF activity at the baseline; however, it is important to note that the nature of GEF activities have shifted since the baseline, and baseline measures of activity may not reflect current activity levels.

The table below provides an overview of the qualitative tools and the number of interviews of each type completed.

*Table 8: Targeted sample sizes for qualitative data collection, by tool or population group*

Method	Source/Respondents
FGD with CEC Members	<ul style="list-style-type: none"> <li>• 10 FGDs total</li> <li>• One FGD conducted per qualitative data collection point</li> <li>• Participants (3-4) selected based on following criteria:               <ul style="list-style-type: none"> <li>○ Must be CEC members for one of the project's intervention schools, including both men and women</li> <li>○ Must represent the same CEC</li> <li>○ Participants must be currently active in the CEC</li> <li>○ Participant should have received coaching or training from the project</li> </ul> </li> </ul>
KII with MOE officials	<ul style="list-style-type: none"> <li>• 6 KIIs total</li> <li>• Regional Education Officers (REOs) in the project areas</li> <li>• Participants selected from list of MOE officials provided by CARE</li> </ul>
FGD with teachers	<ul style="list-style-type: none"> <li>• 10 FGDs total</li> <li>• One FGD conducted per qualitative data collection point</li> <li>• Participants (4-6) selected based on following criteria:               <ul style="list-style-type: none"> <li>○ Must be teachers currently working in this school.</li> <li>○ Must include both male and female teachers, if possible</li> <li>○ Must be teachers who participated in the numeracy training and/or those who received coaching from the project</li> </ul> </li> </ul>
FGD with mothers	<ul style="list-style-type: none"> <li>• 10 FGDs total</li> <li>• One FGD conducted per qualitative data collection point</li> <li>• Participants (4-6) selected based on following criteria:</li> </ul>

<sup>39</sup> Girls with hearing, communication, and cognitive impairments were not included, because conducting a KII with such girls – especially without special accommodations or knowledge (such as knowledge of sign language) on the part of the interviewer – would be difficult for field teams.

<sup>40</sup> Midway through fieldwork, amendments were made to this sampling approach, as some girls were not available for interviewing, could not be located, or did not self-report as having a disability. In the latter case, the majority of the interview questionnaire would have been impossible to complete, as girls who do not self-identify as disabled cannot be asked questions about how their disability impacts their life.

	<ul style="list-style-type: none"> <li>○ Must be mothers of girls enrolled in or recently enrolled in this intervention school</li> <li>○ None can be serving on the CEC or in any education profession (e.g. current or former teachers, school workers)</li> </ul>
KII with girls with disabilities	<ul style="list-style-type: none"> <li>● 10 KIIs total</li> <li>● Participants selected from a list of girls identified at the baseline. <ul style="list-style-type: none"> <li>○ Girls with impaired vision, impaired mobility, or severe anxiety or depression targeted for participation</li> <li>○ Girls with impaired hearing or cognitive functioning not targeted<sup>41</sup></li> </ul> </li> </ul>
Participatory Risk Mapping	<ul style="list-style-type: none"> <li>● <b>4 Risk Mapping Forums total</b> <ul style="list-style-type: none"> <li>○ Participants (4-8 between ages of 12 and 20) selected from Girls Empowerment Forums (GEFs)</li> <li>○ Mapping of areas where children feel comfortable, where they play, etc.</li> </ul> </li> </ul>
Participatory Vignette Exercises	<ul style="list-style-type: none"> <li>● <b>4 Vignette Exercises total</b> <ul style="list-style-type: none"> <li>○ Participants (4-8 between ages of 12 and 20) selected from Girls Empowerment Forums (GEFs)</li> <li>○ Participants presented with the premise of a story about a girl, and asked to conclude the story, with probing about the constraints faced by the girl in the story, how things could have turned out differently, etc.</li> </ul> </li> </ul>

### 3.3. Limitations and Challenges

- **Non-random assignment:**

Non-random assignment to intervention versus comparison sites presents a primary limitation to our ability to make valid causal inferences on the basis of the data collected. The sample design has paired intervention and comparison schools such that they are as balanced as possible in terms of several potentially relevant characteristics. However, as the baseline evaluation showed, intervention and comparison schools were not perfectly balanced at the outset of SOMGEP-T; for instance, intervention schools were more likely to have an established CEC at the baseline, which was a SOMGEP intervention, and it is almost certain that intervention and comparison schools are also imbalanced in terms of other potentially important, but unobserved, factors that may bias analysis. The main implication of this limitation is that, when making inferences on the basis of these data, we cannot be absolutely certain that observed results are a product of program interventions and not at least partly a product of unobserved, systematic, differences between the intervention and comparison groups. We will attempt to mitigate this problem in our analysis using statistical controls in regressions to adjust findings for the influence of observable factors that are significantly different between intervention and comparison groups. However, we can never be certain that we have accounted for all potential confounders, and thus we can never claim that our estimates are completely unbiased.

- **Panel attrition through out-migration:**

---

<sup>41</sup> The use of sign language is very limited in Somalia and virtually unknown in the rural and remote areas where the project is being implemented. Therefore, enumerators will not be able to engage with girls with severe hearing disability who do not know how to read. Similarly, girls with cognitive disabilities may not be able to fully understand the questions posed, raising serious questions about the validity of the answers and reducing the value of the interview.

As with most panel surveys, attrition poses a significant threat to drawing valid inferences. While the prolonged drought that was noted during the baseline report has subsided in many areas, migration – both temporary and permanent – is a common facet of life in Somalia. In some areas, this is exacerbated by continued, localized drought, and by conflict. The design of the baseline sample took into account the possibility of significant attrition due to out-migration, but minimizing attrition remains a goal of data collection at the midline. If the proportion of displaced households exceeds the anticipated attrition rate embedded in the sample size calculation, the project’s ability to assess impact will be compromised. Beyond migration specifically, the midline evaluation made extensive efforts to reduce sample attrition. The evaluation team prescribed a set of formal procedures to be completed before a girl in the baseline learning cohort was replaced by another girl at midline. These procedures included a minimum of two contact attempts at the girl’s household; two contact attempts via telephone (for those households who provided a contact telephone number at the baseline); inquiring at the school with the head teacher, teachers, CEC members and girls in the targeted girl’s grade/age cohort; and inquiring locally at the mosque or other prominent meeting place in the community.

- **Bias from replacement girls**

As noted in the discussion of panel attrition, the sample design laid out at baseline includes a buffer to reduce the risk of attrition. CARE’s M&E staff designed the sample to provide sufficient statistical power to identify project impact even in the face of high levels of panel attrition. Similarly, the midline data collection process includes, as described elsewhere in this report, strict procedures designed to reduce attrition and, by extension, the number of cohort girls who must be replaced because they cannot be re-contacted.

Despite these efforts, a number of girls (159 at the midline, out of 807 total) could not be located and were replaced. Our concern with regard to replacement girls is that they are often not comparable to the girls who they are replacing. Every girl targeted for re-contact was enrolled at the baseline; however, girls we were unable to locate at the midline are less likely to be enrolled in school and more likely to have left the community than the girls who replace them. This is both a function of the types of girls who drop out of the panel – because girls who cannot be re-contacted may be from poorer families, marginalized communities, or may migrate seasonally – and a function of how replacement girls were selected. Replacement girls were selected directly from schools and are, by definition, enrolled in school at the time of the midline. While we account for and eliminate this bias directly in our analysis of transition rates, it may influence other outcomes, insofar as replacement girls come from more stable families or enjoy other advantages relative to cohort girls who cannot be located.

In Section 2.4, we compare replacement and replaced girls to each other directly. In general, this analysis assuages concerns about the two groups of girls being fundamentally incomparable, in that replaced girls are not especially disadvantaged relative to replacement girls. At the same time, we can only assess comparability on those characteristics that were observed at baseline and midline, and unobserved characteristics may distinguish these two groups. Moreover, girls who cannot be located at the midline may have experienced changes in their life between baseline and midline that we cannot capture, because they were not interviewed at midline; these changes may explain, to some degree, our inability to locate them. In short, while the analysis in Section 2.4 does not suggest any specific reason for concern, it cannot rule out entirely the possibility of bias due to the replacement process.

- **Estimating attendance – inaccuracy of school record-keeping:**

At the outset of the baseline evaluation, the evaluation team and CARE staff recognized that collecting attendance data from school records would be challenging. Specifically, it was hypothesized that school records of attendance would be of poor quality, and would consist of either partial or entirely incomplete records. The baseline data was consistent with this hypothesis, and the evaluation team deemphasized attendance data from school records as a result. Instead, the evaluation team relied most heavily on manual headcounts performed by enumerators, supplemented by estimates of attendance derived from the caregiver module of the household survey. The midline evaluation has continued this approach – and provides additional evidence of the unreliability of school record-keeping – triangulating attendance across

multiple sources to provide an overall picture of attendance rates, rather than a precise count of attendance over the previous year.

- **Removal of Outlier Schools**

As noted previously, several schools that were outliers at the baseline have been excluded from the midline sample. More broadly, several additional schools have been excluded for security and other reasons. These sample changes undermine the research design laid out at the baseline, which matched intervention and comparison schools on the basis of their pre-existing characteristics. The pair-matching design laid out at the baseline has significant advantages: pair-matched designs typically show improved statistical power over difference-in-differences designs that do not incorporate such “pre-processing”; in addition, pair-matched designs have the convenience of ensuring pre-intervention balance on at least some observable characteristics. By altering the sample, the design is no longer pair-matched, as some schools are no longer matched to a specific comparison school. More importantly, the set of schools being excluded are non-random – five high-achieving comparison schools have been excluded, which could produce bias in estimates of project outcomes.

The difference-in-differences research design largely mitigates this bias. While pair-matched designs are preferable in many ways, difference-in-differences does not assume balance on pre-intervention characteristics between intervention and comparison groups. Difference-in-differences explicitly accounts for pre-intervention differences between the two groups. During the analysis, we will subset the baseline sample to include only those schools included in both evaluation waves, such that any comparisons from baseline to midline will include only the 63 schools visited at midline. While the loss of a pair-matched design introduces pre-intervention differences between intervention and comparison groups, the difference-in-differences design is robust to this deviation from the planned design.

- **Measuring Transition without OOS Girls**

As described previously, the midline sample of cohort girls includes only those cohort girls who were enrolled in school at the baseline. For the purposes of measuring changes in learning, attendance, and life skills outcomes, this does not present significant problems: using the set of girls who were included in both waves of the evaluation will allow us to draw conclusions about changes over time in this subset of girls.<sup>42</sup> While we are excluding a significant proportion of girls from the analysis, conclusions drawn from the analysis are still valid for the relevant sub-population studied.

In contrast to learning outcomes, transition is more difficult to measure in this manner. Analysing only the subset of girls who were enrolled at the baseline introduces significant bias in baseline estimates of transition; comparing midline transition rates to this biased estimate is unlikely to produce valid conclusions. To avoid this issue, the analysis of transition in this report focuses on a very specific subset of girls at the baseline and midline – those who were enrolled in school one year prior to data collection. For the baseline transition sample, we calculate transition rates among those girls who reported having been enrolled in the year prior to the baseline. For the midline, we calculate transition rates among girls who were enrolled in the year prior to the midline (in essence, this is the full midline sample, since all sampled girls were enrolled one year prior to the midline). The set of girls in these two groups do not overlap perfectly; however, this approach produces the most comparable samples from baseline to endline.

---

<sup>42</sup> We plan to conduct analysis of changes in these outcomes over time in two ways. First, we will analyze data for those girls included in both the baseline and midline (i.e. those girls who were in-school at the baseline *and* who were successfully located at the midline). Second, we will analyze data for those girls included in both waves *or* their replacements (i.e. those girls who were in-school at the baseline, compared to themselves or their replacements at the midline).



### 3.4. Re-Contact Rates and Panel Attrition

As described in the previous discussion of sampling and methodology, the evaluation of SOMGEP-T relies on a panel design, in which the same girls are tracked over time from the baseline through the midline. Although the difference-in-differences approach does not strictly require a panel design – analysis of repeated cross-sections will produce valid inferences under very similar assumptions as a true panel – it is preferable. A panel study ensures that sampling variation over time (i.e. changes in the composition of the intervention or comparison groups from baseline to endline) is not responsible for any results observed. In a repeated cross-sectional design, it is possible for the sample of intervention or comparison girls at the endline to be fundamentally different from the sample drawn at the baseline – if this is true of only intervention or only comparison girls, it would lead to bias in the difference-in-differences model. Likewise, in designs with significant panel attrition and replacement of girls from the original cohort, it is possible for replacements in either intervention or comparison to structurally differ from the girls they are replacing, producing differential trends – i.e. non-parallel trends – that violate the assumptions of the difference-in-differences method.

For the reasons outlined above, limiting panel attrition and ensuring that replacement girls are broadly similar to those cohort girls being replaced are essential goals of the evaluation. This midline evaluation provides an opportunity to study re-contact and replacement systematically. In the discussion that follows, we report the overall re-contact rate, and assess predictors of successful re-contact, to determine whether specific groups of girls were more or less likely to be re-contacted successfully. We also describe the causes of attrition, to the extent that this information could be gathered, and investigate the similarities and differences between replacement girls and the girls they were selected to replace.<sup>43</sup>

#### Re-Contact Rates, Attrition, and Statistical Power

The original midline sample targeted 828 cohort girls. As noted previously, the midline sample differed in two important ways from the baseline: the set of schools was more limited than at the baseline, and only girls who were enrolled in school at the baseline were re-contacted at the midline. The final targeted sample was 824 cohort girls after adjustments were made during fieldwork; of those girls targeted, data was collected for the cohort girl or a replacement in 822 cases. A further 15 girls were removed from the dataset during cleaning, resulting in a sample of 807 cohort girls who were re-contacted or replaced.

Overall, the re-contact rate in the sample was 80.3 percent, with 159 cohort girls of 807 replaced at midline. Although there are theoretical reasons why we might expect re-contact rates to be higher in intervention villages – specifically, CARE’s local footprint, interaction with school officials, and general rapport with the communities in which they work – this was not the case.<sup>44</sup> In intervention areas, 79.6 percent of girls were re-contacted successfully, compared to 81.1 percent of girls in comparison areas ( $p = 0.59$ ). The gap in re-contact rates between intervention and comparison areas does not increase, even in the face of more extensive regression analysis, reported in the next section, that controls for zone, age, and household

---

<sup>43</sup> To fix our terminology in this section and the sections that follow: successful re-contact and attrition are antonyms in this discussion. Our analysis of re-contact rates is limited to those girls who were included in the midline sample, i.e. those girls who we sought to re-contact from the baseline, which is a subset of all baseline cohort girls, as noted in the methodological discussion above. Replacement girls are those who were selected during the midline data collection to replace cohort girls who were not re-contacted successfully (we refer to these latter girls occasionally as “replaced girls”).

<sup>44</sup> Given the short exposure time of communities to the SOMGEP-T interventions at the time of data collection, this may be overstating the potential impact of the project, thus far, on re-contact rates. To the extent that the project improves re-contact rates, these effects are likely dwarfed by exogenous factors that drive migration patterns in the region, including drought, pastoralist livelihoods, and local conflict.



characteristics.<sup>45</sup> Based on these results, we are confident that differential attrition between intervention and comparison areas is not a concern at this stage of the evaluation.

The SOMGEP-T sampling strategy was developed with a high level of panel attrition in mind. In the project's MEL Framework, the sampling calculations included a buffer for up to 40 percent attrition. At this stage, the evaluation is approximately within the expected range of attrition, at 19.7 percent from baseline to midline. A more precise and conservative statement of the attrition rate from baseline to midline would include those girls whose data was removed during the cleaning stage. As noted above, 824 girls were targeted in the final midline sample, and 648 were successfully re-contacted, with the remainder either replaced or dropped from the sample entirely during the cleaning stage. This results in a more conservative re-contact rate of 78.6 percent or a corresponding attrition rate of 21.4 percent.

An important open question concerns how attrition rates will differ from the 2018 midline to the 2019 second midline and from the 2019 midline to endline. Two considerations make it difficult to draw firm conclusions about future attrition rates. First, because the midline only sought to re-contact girls who were in school at the baseline, attrition rates may be higher among the full sample of cohort girls and go beyond the assumption of 40 percent. This would occur if OOS girls were less likely to be successfully re-contacted later, if they are more likely to migrate, get married, or have less stable households than in-school girls. However, because all girls were recruited through their households and completed the household survey, the evaluation team has the same detailed contact information for all girls, whether enrolled or out-of-school at baseline. For this reason, we do not expect OOS girls to have significantly higher attrition rates than in-school girls, all else equal.

Second, attrition rates are likely to differ over time, though it is not possible to make specific predictions in this regard. On one hand, attrition rates may decline over time, because the girls most likely to drop out of the sample have already dropped out between baseline and the first midline; in that case, the sample becomes an increasingly self-selected group of girls who are more and more likely to remain in the panel over time. It is also possible that the project itself will improve transition rates, reducing panel attrition over time, at least in intervention areas. On the other hand, a number of unknown factors will influence attrition rates over the next two years, including rainfall patterns, migration, and conflict dynamics, and it is not possible to predict the direction or magnitude of their effects on attrition rates.

Ultimately, the sample size from round to round is preserved through the replacement process. That is, the sample size for analysis of midline to endline changes will be very similar to the sample for analysis of baseline to midline changes, because the panel was "refreshed" during this midline round and will also be refreshed at the time of the second midline. At the same time, the "true panel" sample is diminished over time as additional girls from the baseline drop out of the panel. If we make the following assumptions regarding attrition rates:

- Identical attrition rates among out-of-school and in-school girls;
- An identical attrition rate from midline to endline as was observed from baseline to midline;
- The re-inclusion of all schools removed from the midline sample, except those five schools that were baseline outliers,

then the true panel available for analysis from baseline to endline would be reduced to 1005 cohort girls, or 57.7 percent of the baseline sample size. This reduction comes about due to traditional individual-level attrition and the blanket removal of five outlier schools from the sample. In total, this estimated sample size falls just above the target established by CARE at the project's outset: in line with the FM's guidance, the project set a target of 502 girls each, after attrition, in intervention and comparison areas (n = 1004 total),

---

<sup>45</sup> In a bivariate t-test, comparison areas had re-contact rates 1.5 percentage points higher than intervention areas; in the regression models reported in the next section, this gap fluctuated from 1.2 to 1.6 points, and never approached statistical significance at any conventional level.

targets that took into account clustering and design effect, as well as the goal effect size for learning. While we have some concern about the assumptions that underlie these sample size calculations, the current attrition trend suggests that the target will be met, or very nearly met, at the endline.<sup>46</sup> At the midline, panel attrition was high, but manageable; additional efforts will need to be made at the second midline and the endline to increase successful re-contacts as much as possible.

### Predictors of Successful Re-Contact

In the vast majority of cases where girls were not re-contacted successfully, two outcomes occurred. The most common outcome was that the household had moved out of the community. In 49.4 percent of unsuccessful re-contacts, the household had left the community. The second most common outcome was that the household remained in the community, but the girl had left the household; this disposition applied to 28.1 percent of unsuccessful re-contacts. The remaining unsuccessful re-contacts were primarily due to idiosyncratic circumstances or cases where enumerators did not clearly specify the reason that replacement was needed.<sup>47</sup>

Re-contact rates varied somewhat by age, the latter of which is shown in the table below.

*Table 9: Re-contact rates by age and grade*

Age	Intervention Area Re-Contact Rate	Comparison Area Re-Contract Rate	Overall Re-Contact Rate
10	94.2%	90.0%	92.3%
11	77.4%	88.5%	82.5%
12	86.1%	83.3%	84.8%
13	81.9%	79.0%	80.6%
14	71.2%	70.0%	70.1%
15	71.8%	90.2%	81.3%
16	70.0%	68.8%	69.6%
17-19	61.9%	57.9%	61.0%
Total	79.6%	81.1%	80.3%
Grade			Re-Contact Rate
3 (3)	77.7%	78.3%	78.0%
3 (4)	83.3%	89.3%	85.9%
4 (5)	76.6%	86.4%	81.4%
5 (6)	80.9%	69.7%	75.8%

As the table shows, age is strongly correlated with re-contact rates, a fact that is supported by additional analysis below. Among girls aged 10-12 years, 86.6 percent were re-contacted; among girls 13-15 years, that rate dropped to 76.8 percent; and among girls 16 years and older, just 65.5 percent were successfully

<sup>46</sup> The FM's guidance suggests an intra-cluster correlation (ICC) of 0.1, from which design effect assumptions are derived and used to adjust the target sample size. However, at baseline, intra-cluster correlation among non-outlier schools for numeracy was 0.19 and for literacy was 0.15. Applying these ICC values to the sample size targets – and making no other adjustments to the assumptions used – would inflate them from 1004 girls to between 1104 and 1264 girls. Under this updated assumption regarding ICC, the current attrition rate is too high to meet the necessary sample size for the project's target statistical power. However, it is important to note that this is not a function of attrition rates, per se, but a function of the assumptions underlying the power calculations and, under the FM's guidance for sample size calculations, the project is approximately on pace to meet the sample size targets after attrition.

<sup>47</sup> Of the 160 unsuccessful re-contacts, just 3.1 percent of the girls were still part of a household in the village but were unavailable for an interview at the time of our fieldwork visits. Even fewer girls – 0.6 percent, or 1 girl out of 160 – were part of households that refused to participate.

re-contacted. Given that older girls are more likely to migrate due to marriage or to seek work in urban areas, this trend is not altogether surprising. The overall effect of the relationship between age and re-contact rates is that "true panel" sample has become slightly younger than it otherwise would have been, if re-contact was successful in all cases. The mean age of re-contacted girls at midline was 13.63 years, compared to 14.57 years for girls who were not successfully re-contacted and required replacement ( $p < 0.00$ ). This difference results in only a small change to the mean age of the true panel sample, however: in the case of 100 percent successful re-contact, the mean age of the sample analysed at midline would have been 13.82 years; in practice, the mean age of the re-contacted girls at midline was 13.63 years.<sup>48</sup>

In contrast, grade is a less straightforward predictor of re-contact rates. The lowest re-contact rates are in grades 5 (6) and 3 (3), with notably higher rates in grades 3 (4) and 4 (5). While age is a strong predictor of re-contact rates, grade may be a noisier predictor because age and grade are not perfectly correlated in the Somali context. Nonetheless, it is unusual that higher grade levels are not associated with monotonically decreasing re-contact rates; it is possible that the grade 3 (3) results are simply an artefact of sampling variance, but we cannot draw firm conclusions based on the data available.

To better assess the multiple factors that could influence attrition and re-contact, we estimated a series of regression models predicting successful re-contact on the basis of a girls' baseline characteristics and those of her household. Our first model incorporated dichotomous (dummy) variables for intervention status and zone of residence at baseline, and an additional variable capturing each girl's age at baseline. The results of this regression are presented in the left panel of the figure below: each blue dot is the regression coefficient for a given variable, and the bars represent the 95 percent confidence interval around that coefficient. The dark vertical line represents a null effect – confidence interval bars that cross the vertical line at zero are statistically insignificant.

In this regression model, we incorporate variables that indicate whether a household migrates seasonally, has a female head of household, owns a mobile phone, and solely owns land.<sup>49</sup> We also estimated models that incorporated baseline learning scores, theorizing that lower-performing girls may be less likely to stay in school and more likely to migrate, get married, or otherwise leave the community. Numeracy and literacy scores, alone or combined, do not appear to influence re-contact rates.<sup>50</sup>

Notably, indicators of household wealth or economic security are not good predictors of successful re-contact. In additional regressions, available upon request, we utilized alternative economic indicators – including the quality of a household's roof, the frequency with which they go to bed hungry, the frequency with which they lack water for household needs, the frequency with which they lack a cash income, and whether they rely on charity to meet their basic household needs – to study the role of household wealth in re-contact rates. None of these alternative measures were statistically significant, and the direction of their effects varied widely; for instance, a girl living in a household that self-reported being unable to meet its basic needs without charity was 4.4 percentage points *more likely* to be successfully re-contacted at the midline,

---

<sup>48</sup> As shown in Table 9, re-contact rates varied by age across intervention and comparison areas. However, in the aggregate, there is no observable difference between intervention and comparison areas in terms of the effect of age on attrition or re-contact. In other words, the general trend of higher re-contact rates among younger girls – and the corresponding reduction in the mean age of the "true panel" sample – is consistent between intervention and comparison areas.

<sup>49</sup> Seasonal migration is coded based on self-reports at the baseline, with 3.8 percent of household heads stating that their household occasionally migrate as a result of seasonal changes. Land ownership was also based on self-reports: households could indicate whether they own land solely, own land jointly with other households, own some land solely and some jointly, or do not own land. We consider the sole ownership of land as a metric of household economic status, even if the household *also* owns land jointly with other families.

<sup>50</sup> A 10-point increase (on a 100-point scale) in numeracy at baseline is associated with a 0.4 percent decline in the likelihood of being successfully re-contacted, while a 10-point increase in Somali literacy is associated with a 0.7 percent increase in the same likelihood. Neither effect is statistically significant at any conventional level.

though this result was not statistically significant.<sup>51</sup> These findings suggest that attrition is not driven, first and foremost, by household poverty, or that poverty interacts with – or is otherwise influenced by – other factors and household characteristics that shape re-contact rates.<sup>52</sup>

## Replaced and Replacement Girls

Data collection teams at the midline were provided detailed instructions for locating cohort girls, as well as for replacing cohort girls who could not be located. Enumerators were instructed to select replacement girls from the same school and grade, if such a girl were available to interview. In practice, some schools have too few students to allow such like-for-like replacement, and many deviations occur in the data. Even replacements that followed these criteria, moreover, invite concern about the comparability of replacement girls to those they replaced, as replacement was based on sharing a grade level and no other observed or unobserved characteristics.

Usefully, the richly detailed data on cohort girls at the baseline and their replacements at the midline allow us to compare the two sets of girls (n = 160 in each group). To assess their comparability, we tested the differences between replaced and replacement girls in terms of age, grade, share in female-headed households, economic status indicators, literacy and numeracy scores, self-reported attendance in the year of data collection, and control over their own schooling decisions. The results of these t-tests are provided in the table below.

Table 10: Comparison of baseline cohort girls and their replacements

Indicator	Baseline Cohort Girls	Replacement Cohort Girls	Difference in Means (p-value)
Age at midline*	14.6 years	13.6 years	1.02 (.00)
Grade	4.3	4.8	0.5 (.00)
Female head of household	47.5%	50.6%	-3.1% (.58)
Improved roof	67.9%	59.1%	8.8% (.10)
HH owns land solely	69.4%	64.7%	4.8% (.39)
HH goes to sleep hungry many/most days	8.1%	7.6%	0.5% (.87)
HH lacks clean water many/most days	26.9%	28.9%	-2.1% (.68)
HH lacks cash income most days	19.5%	22.8%	-3.3% (.47)
HH owns mobile phone	96.3%	90.0%	6.3% (.03)

\*Age for baseline cohort girls was extrapolated to the midline by simply adding one year, as midline data collection occurred almost exactly one year following baseline data collection.

As the results in the table show, there are three important demographic differences between replaced and replacement girls. The first is their age: replacement girls are younger than replaced girls *would have been at the midline*, by approximately one year on average. While the modal replacement girl was the same age as

<sup>51</sup> In fact, as their household's ability to meet their own needs increased, girls became less and less likely to be successfully re-contacted – the lowest re-contact rates occurred among the two highest self-sufficiency (i.e. able to meet their own needs and purchase non-essential goods occasionally) categories.

<sup>52</sup> While the results reported here are based on linear (OLS) regression models, we also tested the robustness of the results to the use of logistic regression, in line with the binary nature of the outcome variable, successful re-contact. We prefer the linear models for the ease of their interpretation, but the substantive conclusions regarding the correlation – or lack thereof – between attrition, age, zone of residence, and household economic status are not affected by the specification choice.

the replaced girl *would have been* (n = 39 or 24.4 percent of all cases), other replacement girls tended to be younger than those girls they replaced by one (16.3 percent of cases) or two (18.1 percent of cases) years.<sup>53</sup>

Second, replacement girls tended to be slightly ahead of replaced girls in terms of grade level. Importantly, this measure compares replacement girls at midline to replaced girls at baseline, so this gap is likely overstated; because we cannot know how many replaced girls advanced grade levels between baseline and midline, it is not possible to test this difference precisely.

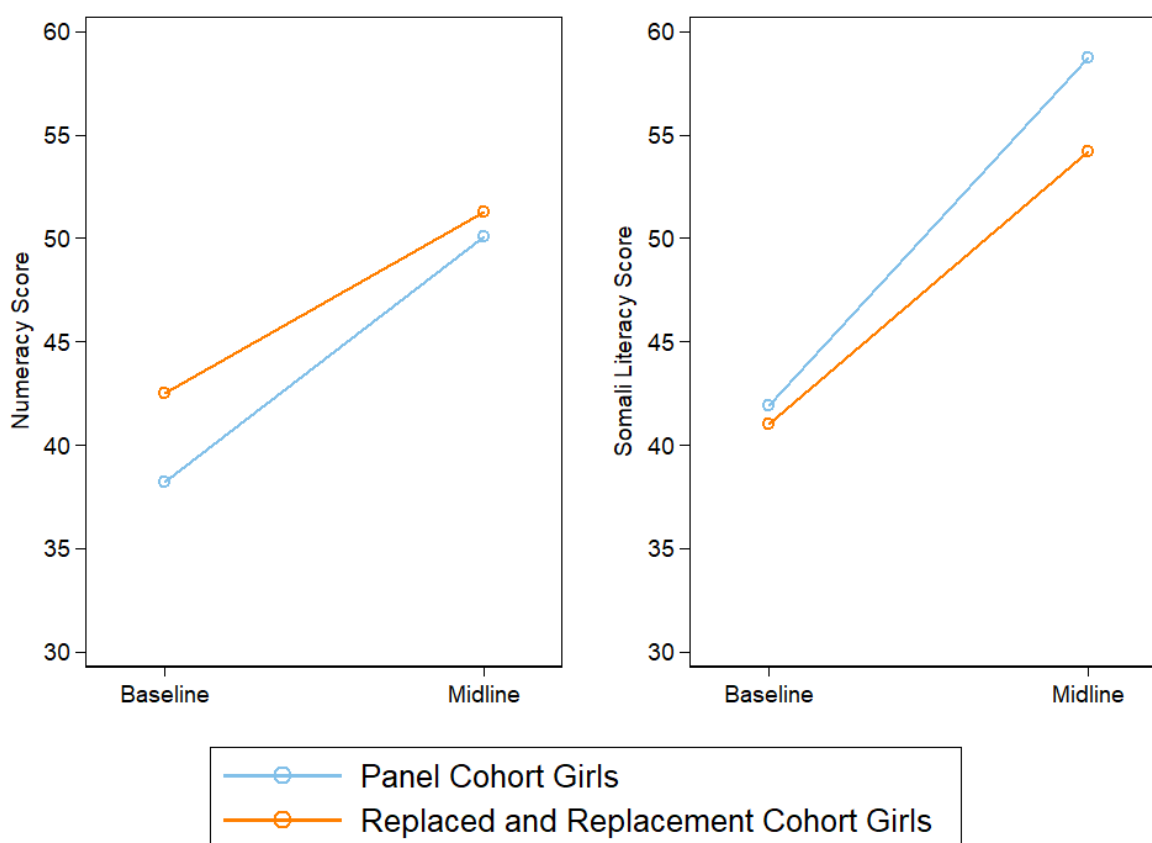
Third, replacement girls tended to be from households that were disadvantaged, relative to those girls they replaced. In terms of economic deprivation – going to sleep hungry at night, lacking clean water for household use, and lacking a cash income – replacement and replaced girls' households were comparable. However, replacement girls' households were less likely to have an improved roof, and were much less likely to own a mobile phone than the households of girls they replaced.

Beyond demographic factors, how do replaced and replacement girls compare? Replacement girls score markedly better on both numeracy and literacy tests than the girls they replaced at baseline, scoring 8.8 points higher on numeracy and 13.2 points higher on Somali literacy. However, it is important to note that a similar – in fact, slightly more dramatic – upward trend was seen among cohort girls who were successfully re-contacted at the midline. Among girls who constitute a true panel, numeracy scores rose 11.9 points and Somali literacy scores rose 16.8 points. The figure below plots these trends, in panel cohort girls and replacement/replaced cohort girls, from baseline to midline (numeracy in the left panel and Somali literacy in the right panel). As these results show, trends were dramatically upward in both groups, and replacement girls performed – relative to those girls they replaced from baseline – within a reasonable range. Based on these results, there is no reason to expect that replaced and replacement girls differ systematically in terms of learning scores, though we cannot rule out this possibility based on the available data.

---

<sup>53</sup> This difference interacts with the general trend, noted above, that re-contact rates were higher among younger girls. As discussed above, the "true panel" sample became slightly younger as a result of differential attrition by age; combined with younger replacements, the overall sample (i.e. including both panel and replacement girls) has become slightly younger than it otherwise would have been at the midline, from an expected mean age of 13.82 years to a realized mean age of 13.6 years.

Figure 1: Learning scores for “panel” girls versus replaced and replacement girls



On other indicators, replacement girls were either similar to the girls they replaced or slightly disadvantaged. Attendance rates, as reported by their primary caregivers, were very slightly lower among replacement girls – 88.1 percent of replacement girls were reported to attend school most days, compared to 90.6 percent of replaced girls at the baseline. Replacement girls were also slightly less likely to state that they had control over their own marital decisions. Note, however, that the differences cited regarding attendance rates and control over marital decisions were not statistically significant at any conventional level. One area where replaced and replacement girls differed was in control over the decision to go to school or remain in school: 37.6 percent of replaced girls indicated they had sole control over this decision at the baseline, compared to just 25.5 percent of replacement girls at the midline.<sup>54</sup>

In general, there do not appear to be systematic differences between cohort girls from the baseline evaluation who were replaced and those girls who replaced them. While replacement girls are somewhat younger than their counterparts, they are also more economically disadvantaged. And, while their learning scores are considerably higher than the girls they replaced, this is consistent with higher scores across the board at the midline, even for girls who were successfully re-contacted and completed both baseline and

<sup>54</sup> This gap does not appear to stem from differences in age. Girls who are replaced (not successfully re-contacted) are just under one year older, on average, than their cohort peers who were successfully re-contacted and are, in turn, very slightly more likely to have influence over their own schooling decisions ( $p = 0.72$ ). However, girls who are replaced are *not* systematically older than the girls who replace them *at the time of data collection* – the average age of replaced girls was 13.57 years during data collection at baseline, compared to 13.54 years for replacement girls during midline data collection – but are nonetheless much more likely to claim influence over their own schooling decisions ( $p = 0.05$ ).



midline learning assessments. The available data does not allow us to draw absolute conclusions regarding the comparability of replaced and replacement girls; indeed, unimpeachable conclusions on this question are impossible in most circumstances, since we never observe the changes that occur in replaced baseline girls between baseline and midline. However, given the available data, we do not find any clear reason to question the comparability of replaced and replacement girls.

## 4. Context, Educational Marginalisation and Intersection between Barriers and Characteristics

### 4.1. Educational marginalisation

The following tables present the breakdown of sample girls across the baseline and midline evaluations who are affected by characteristics or barriers related to their educational marginalization. The figures below only include the panel of girls interviewed both at the baseline and midline evaluation. Thus, the comparison between rounds presents information regarding the changes in the distribution of characteristics and barriers for the same individuals and are not reflective of changes in sample composition.

#### Characteristics

The table below presents the proportion of girls with some of the main characteristics related to their marginalization. SOMGEP-T defines educational marginalization characteristics as including poverty, family conditions, language obstacles, and illiteracy among caretakers.

There were significantly fewer primary caregivers in the midline who said it was difficult to afford schooling for their girl than did so in the baseline. Also, in the midline, 7.6 percent of primary caregivers in comparison areas said that they have difficulty in affording their girls' school versus 18.1 percent from the baseline, and 12.9 percent of primary caregivers in intervention areas said that they have such difficulty in the midline compared to 21.1 percent from the baseline. Difference-in-difference analysis does not show that households in intervention areas did significantly better than those in comparison areas. This improvement in livelihood among respondents may instead be related to the recovery from a past drought. Indeed, baseline data was collected after two years of drought that put a strain on the communities, while the midline data has been collected after a better than average rainy season.

The lack of education among households and primary caregivers has decreased significantly as well.<sup>55</sup> These changes may be attributed to caregivers attending non-formal education courses offered in the area by CARE or other organisations. The changes may also be due to the respondents understanding the question differently between baseline and midline, e.g. caregivers did not consider non-formal education as education in the baseline but did so in the midline. Thus, this change in proportions from baseline to midline should be carefully considered.

However, the rest of the characteristics affecting girls' marginalisation have remained constant throughout the evaluated period of time. No statistically significant differences can be observed across intervention and comparison areas between baseline and midline rounds.

*Table 11: Girls' characteristics*

	Baseline		Midline		
	Intervention	Comparison	Intervention	Comparison	ALP

<sup>55</sup> Significance tests on the correlation between rounds and education of heads of household and of primary caregivers provide results of P-values of 0.00 for both variables across all evaluated areas.



Sample breakdown (Girls)					
Single orphan	38 (11.1%)	30 (9.8%)	47 (13.7%)	40 (13.1%)	53 (14.5%)
Double orphan	1 (0.3%)	0 (0%)	0 (0%)*	0 (0%)	0 (0%)
Living without both parents	21 (6.1%)	18 (5.9%)	32 (9.3%)	25 (8.2%)	64 (17.5%)
Living in female headed household	146 (42.6%)	136 (44.6%)	156 (45.5%)	137 (44.9%)	151 (41.4%)
Married	3 (0.9%)	3 (1%)	4 (1.2%)	6 (2%)	70 (19.2%)
Mother under 18	1 (0.3%)	0 (0%)	1 (0.3%)	4 (1.3%)	8 (2.2%)
Mother under 16	1 (0.3%)	0 (0%)	0 (0%)*	3 (1%)*	2 (0.5%)
Difficult to afford for girl to go to school <sup>56</sup>	70 (21.1%)	54 (18.1%)	42 (12.9%)*	21 (7.6%)*	0 (0%)
Home uses poor roofing material*	111 (32.4%)	115 (37.7%)	126 (36.7%)	122 (40%)	154 (42.2%)
Gone to sleep hungry for many days in past year <sup>57</sup>	25 (7.3%)	26 (8.6%)	31 (9.1%)	39 (12.9%)	31 (8.5%)
HoH has no education <sup>58</sup>	217 (67%)	181 (64%)	162 (49.1%)*	152 (51.2%)*	137 (37.5%)
Primary caregiver has no education	232 (67.6%)	229 (75.1%)	176 (51.3%)*	182 (59.7%)*	167 (45.8%)
Total girls	343 (100%)	305 (100%)	343 (100%)	305 (100%)	365 (100%)

\*Note, an asterisk in the following tables indicates results that are statistically significant at the 95% confidence level (or higher) in a regression with cluster-robust standard errors.

The table below illustrates the evaluation sample by disability and type of disabilities of the cohort girls across the baseline and midline for intervention and comparison areas. This table presents the proportion of girls who had a severe level of disability in which they said they could not perform a given task at all due to their disability, said that they had “a lot of difficulty” in doing so, or were affected by it at least monthly.

As reported physical, cognitive and communication impairments are generally rarer in the midline than in the baseline evaluation, almost all the girls who were evaluated to have any disability at all are affected by a mental health impairment. It should be noted that while mental health impairments were more common in the midline, the increased proportion of girls affected by the impairments are largely due to a broader population of girls who were sampled in the midline. The proportion of girls with disabilities has increased to 14.3 percent in the intervention areas and 14.4 percent in the comparison areas as compared to the baseline’s 7.6 percent and 5.9 percent per respective area.<sup>59</sup> On the other hand, girls with cognitive disabilities are significantly less common in intervention areas in the midline than they were in the baseline, and girls with hearing disabilities are also less common in intervention as well as comparison areas in the midline. Yet, the difference in both cases is less than one percentage point, which suggests the non-substantial character of

<sup>56</sup> Percentages here are calculated from 332 girls in intervention/baseline, 298 comparison/baseline, 325 intervention/midline, and 278 comparison/midline.

<sup>57</sup> Percentages here are calculated from 342 girls in intervention/baseline, 303 comparison/baseline, 339 intervention/midline, and 302 comparison/midline.

<sup>58</sup> Percentages here are calculated from 324 girls in intervention/baseline, 283 comparison/baseline, 330 intervention/midline, and 297 comparison/midline.

<sup>59</sup> Significant tests on the correlation between rounds and girls with any disabilities provide a P-value of 0.043 for intervention areas and a P-value of 0.003 for comparison areas.

such changes although statistically significant, as most of these physical disabilities disappeared from the midline sample.

As mentioned above, the increased proportion of girls with any disability is mainly due to the different sample of caregivers answering the questions about mental health in the midline evaluation. The questions to the caregivers used to evaluate whether a girl had a mental health impairment were “How often does GIRL seem very anxious, nervous or worried?” and “How often does GIRL seem very sad or depressed?” With CARE, it was decided that if the caregiver said that the girl experienced such feelings daily, weekly, or monthly, then the girl was evaluated to have a mental health impairment. However, the baseline evaluation only included the responses of primary caregivers whose girls were under 12 years of age, while during the midline evaluation, the same questions were also submitted to the caregivers whose girls were 12 years of age or older as well. Since all caregivers were asked these questions about the mental health of their girls in the midline instead of a sub-population as in the baseline, a greater proportion of girls was expected to be evaluated to have mental health impairments in the midline than in the baseline, and indeed we find that to be true as shown in the table below.

*Table 12: Evaluation sample breakdown of girls (by aggregated disabilities)*

	Baseline		Midline		
	Intervention	Comparison	Intervention	Comparison	ALP
<b>Sample breakdown (Girls)</b>					
Girls with any disability	26 (7.6%)	18 (5.9%)	49 (14.3%)	44 (14.4%)	29 (7.9%)
Vision impairment	2 (0.6%)	1 (0.3%)	1 (0.3%)	2 (0.7%)	1 (0.3%)
Hearing impairment	3 (0.9%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Mobility impairment	3 (0.9%)	0 (0%)	1 (0.3%)	0 (0%)	1 (0.3%)
Cognitive impairment	2 (0.6%)	1 (0.3%)	0 (0%)	0 (0%)	0 (0%)
Self-care impairment	2 (0.6%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Communication impairment	3 (0.9%)	1 (0.3%)	0 (0%)	1 (0.3%)	0 (0%)
Mental health impairment	19 (5.5%)	16 (5.2%)	48 (14%)	42 (13.8%)	28 (7.7%)
Total girls	343 (100%)	305 (100%)	343 (100%)	305 (100%)	365 (100%)

Using only the comparable sample of 241 girls whose caregivers answered questions about disabilities in *both* the baseline and midline evaluation, we find that the prevalence of mental health impairments has declined among the girls in the comparable sample in both the intervention and comparison groups. There was not statistically significant change in the proportion of comparison girls with mental health impairments between baseline and midline. However, among intervention girls, there was a statistically significant decrease, at the 90 percent confidence level, in mental impairments, as shown in the table below. The proportion of intervention girls with mental health impairments was halved from 17.1 percent in the baseline to 8.5 percent in the midline. Although difference-in-difference analysis does not suggest a significant difference in the decline of mental health impairments of girls in the intervention areas relative to those in comparison areas, this decline in the prevalence of mental health impairments in the intervention group relative to the comparison group suggests that the intervention may be having a positive impact on girls’ mental health.

*Table 13: Evaluation sample breakdown of mental health impairments of only comparable girls*

	Baseline		Midline	
	Intervention	Comparison	Intervention	Comparison
<b>Sample breakdown (Girls)</b>				
Mental health impairment	22 (17.1%)	18 (16.1%)	11 (8.5%)	16 (14.3%)
Total girls	129 (100%)	112 (100%)	129 (100%)	112 (100%)

The following table presents the evaluation sample by the disaggregated disabilities of girls across intervention and comparison schools in the baseline and midline. As it disaggregates the variables presented in the previous table, the table presents the proportion of girls who had a high level of disability in which their caregivers 1) said they could not perform a given task at all due to their disability, 2) said that they had “a lot of difficulty” in doing so, or 3) were affected by it at least monthly. The data show no significant changes from the baseline evaluation overall, but a general pattern of positive correlation between midline and disabilities can be observed. Most of these disabilities have increased even if insubstantially, and the most common disabilities concerning vision and serious illnesses have affected more girls by two to three percentage points in the midline.

*Table 14: Evaluation sample breakdown of girls (by disaggregated disabilities)*

	Baseline		Midline		
	Intervention	Comparison	Intervention	Comparison	ALP
<b>Sample breakdown (Girls)</b>					
Difficulty seeing even wearing glasses	2 (0.6%)	1 (0.3%)	1 (0.3%)	2 (0.7%)	1 (0.3%)
Wears glasses	0 (0%)	0 (0%)	4 (1.2%)	5 (1.6%)	6 (1.6%)
Difficulty hearing sounds even with hearing aid	2 (0.6%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Difficulty walking or climbing steps	3 (0.9%)	0 (0%)	1 (0.3%)	0 (0%)	1 (0.3%)
Difficulty remembering things or concentrating	1 (0.3%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Difficulty with self-care such as washing all over or dressing	1 (0.3%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Difficulty communicating	1 (0.3%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Had any serious illnesses	32 (9.3%)	21 (6.9%)	43 (12.5%)	32 (10.5%)	53 (14.5%)
Difficulty making friends	2 (0.6%)	1 (0.3%)	0 (0%)	1 (0.3%)	0 (0%)
Seems very anxious, nervous or worried	19 (5.5%)	12 (3.9%)	40 (11.7%)	36 (11.8%)	20 (5.5%)
Seems very sad or depressed	11 (3.2%)	14 (4.6%)	33 (9.6%)	25 (8.2%)	19 (5.2%)
Total girls	343 (100%)	305 (100%)	343 (100%)	305 (100%)	365 (100%)

## Barriers

The table below shows the proportion of girls in the sample who face potential barriers to learning and transition in the domains of safety, parental/caregiver support, attendance, school facilities, and teachers across comparison and intervention areas from the baseline to the midline evaluation. There have been

some improvements in teaching quality and education facilities from the baseline evaluation, although many of these changes have occurred in comparison schools.

There have been statistically significant changes in whether girls feel safe at school as well as their sense of choice in attending school. In addition, 3.7 and 2.5 percent of girls respectively in the intervention and comparison areas do not feel safe at school in the midline, while in the baseline 3.2 percent and 9 percent of the girls, respectively, said the same. The data also show a significant increase in the proportion of girls in comparison areas who feel they have no choice in whether to attend school.<sup>60</sup>

Regarding the conditions of schools and teaching, there have been significant positive changes in the seating available and drinkable water facilities: 30.5 percent of girls in comparison schools said their school does not have seats for all student in the baseline and in the midline 15.8 percent did so, and the proportion of girls in intervention schools not using drinking water systems dropped from 19.7 percent in the baseline to 10.5 percent in the midline evaluation.

Teachers' absence has also significantly decreased in intervention schools, compared to comparison ones,<sup>61</sup> as 17.2 percent of girls in intervention areas in the midline said that their teachers are often absent from class versus 35 percent in the baseline, while 29.2 percent of girls in comparison areas reported teachers' absence in the midline versus 40.1 percent of the baseline. Findings on teacher absenteeism will be discussed in more depth in the school governance and management section of the report, but it should be noted here that this finding is likely related to improvements in record-keeping and overall monitoring of teacher attendance by CECs.

The table below also reveals that in both intervention and comparison areas students reported more frequently being afraid of their teachers:<sup>62</sup> in the midline, 75.9 percent of girls in intervention areas and 82.7 percent of girls in comparison areas said they were afraid of their teacher compared to the 57.7 percent and 58.3 percent, respectively, who did so in the baseline. In addition, there has been an overall increase in girls who said their teacher uses punishment or discipline when students get lessons wrong, from 76.8 percent among girls in intervention areas at baseline to 80.2 percent in the midline, and particularly among girls in comparison schools who were significantly more likely to say that their teacher uses punishment and discipline in the midline, 84.9 percent, than in the baseline, 79 percent. These results may be related to the success of the program and campaigns that may have also reached comparison areas, since disclosure can be the first positive sign of addressing the right issues and now students are starting to openly report problems more often.

*Table 15: Potential barriers to learning and transition*

	Baseline		Midline		
	Intervention	Comparison	Intervention	Comparison	ALP
<b>Sample breakdown (Girls)</b>					
<b>Home – community</b>					
<b>Safety</b>					
Doesn't feel safe travelling to/from school	10 (3.2%)	24 (9%)	12 (3.7%)	7 (2.5%)*	11 (3%)
Girl travels more than 30 minutes to school	24 (7.7%)	11 (4.1%)	17 (5.2%)	18 (6.5%)	0 (0%)

<sup>60</sup> The change between rounds for comparison areas is indeed significant with P-values of 0.03 given linear regression analysis.

<sup>61</sup> The finding is significant with a P-value of 0.05 using a cluster-robust logistic regression.

<sup>62</sup> This change between rounds is indeed significant with P-values of 0.00 for both areas given linear regression analysis.

<b>Parental/caregiver support</b>					
Doesn't get support to stay in school and do well	19 (6.1%)	3 (1.1%)	6 (1.9%)*	8 (2.9%)	15 (4.1%)
Girl has no choice in whether to attend school	245 (86.3%)	220 (85.6%)	279 (86.1%)	261 (93.5%)*	225 (61.6%)
<b>School level</b>					
<b>Attendance</b>					
Attends school half the time	4 (1.2%)	6 (2%)	6 (1.8%)	4 (1.4%)	0 (0%)
Attends school less than half time	3 (0.9%)	2 (0.7%)	1 (0.3%)	8 (2.9%)*	0 (0%)
Doesn't feel safe at school	15 (4.9%)	17 (6.4%)	7 (2.2%)	4 (1.4%)*	0 (0%)
<b>School facilities</b>					
No seats for all students	65 (20.9%)	81 (30.5%)	43 (13.3%)*	44 (15.8%)*	49 (13.4%)
Doesn't use drinking water facilities	61 (19.7%)	85 (32%)	34 (10.5%)*	66 (23.7%)	69 (18.9%)
Doesn't use toilet at school	70 (22.6%)	80 (30.1%)	58 (17.9%)	80 (28.8%)	62 (17%)
No computers in class	276 (89.6%)	252 (95.1%)	309 (95.7%)*	259 (93.2%)	317 (86.8%)
Cannot use books or other learning materials at school	62 (19.9%)	67 (25.1%)	65 (20.1%)	79 (28.3%)	89 (24.4%)
<b>Teachers</b>					
Disagrees teachers make them feel welcome	16 (5.1%)	20 (7.5%)	13 (4%)	16 (5.8%)	9 (2.5%)
Agrees teachers treat boys and girls differently in the classroom	127 (41.2%)	110 (41.4%)	137 (43.1%)	125 (46.5%)	102 (27.9%)
Agrees teachers often absent from class	108 (35%)	107 (40.1%)	53 (17.2%)*	77 (29.2%)*	57 (15.6%)
Afraid of teacher	179 (57.7%)	155 (58.3%)	246 (75.9%)*	230 (82.7%)*	200 (54.8%)
Uncomfortable asking teachers question	17 (5.5%)	13 (4.9%)	5 (1.5%)*	7 (2.5%)	8 (2.2%)
Teacher punishes/disciplines when students get lesson wrong	239 (76.8%)	211 (79%)	260 (80.2%)	237 (84.9%)	170 (46.6%)
Physical punishment witnessed last week	87 (63%)	70 (49.6%)	96 (54.5%)	86 (51.2%)	21 (5.8%)
Caregiver rates quality of teaching as poor	11 (3.3%)	14 (4.8%)	5 (1.6%)	8 (2.9%)	0 (0%)

## Migration of Girls and Boys

In order to understand the level of migration in cohort girls' households, heads of households were asked how many girls and boys aged 11 to 21 years of age of their household migrated away in the past 12 months, why they left, and where they left to.

From baseline to midline, the households in the panel data have experienced more migration of girls across both intervention and comparison groups, as shown by the percent of households with at least one migrant girl and the absolute total number of migrant girls from cohort households in the table below, though only the increase in households with a migrant girl is statistically significant. The households of ALP girls were significantly more likely to have had a migrant girl in the past 12 months. Almost all of these girls left to other Somali villages. In the baseline, most of the migrant girls moved to go to boarding school or to stay with other family members, as 40 percent of migrant girls in intervention areas and 31.1 percent in comparison areas went to boarding school and 45.7 percent in intervention areas and 42.2 percent in comparison areas left to live with other family. In the midline, it appears the increase in migration of household girls is due in part to aging, as seen in the higher proportion of girls who left to work and to get married among girls in both intervention and comparison areas.

In a similar fashion, within the panel, we observed a significant increase in the percent of households with migrant boys from baseline to midline as shown in the table below. The vast majority of these boys left to another Somali village. Regarding their reasons to migrate, in the baseline the majority left to boarding school or to stay with other family members. Compared to the baseline, more boys in the midline have moved to stay with other family members or to work somewhere else: 8.3 percent in intervention areas and 11.9 percent in comparison areas left for work-related reasons in the midline. Like the girls, more migrant boys are now attending school versus to the baseline: 36.1 percent in intervention areas and 31.7 percent in comparison areas attend school while in the baseline respective 13.3 and 12.8 percent did.

Table 16: Migration by Girls and Boys (age 11-21) in Cohort Girl's Household

	Baseline		Midline		
	Intervention	Comparison	Intervention	Comparison	ALP
<b>Sample breakdown (Girls)</b>					
Household with a migrant girl (% of total)	30 (8.8%)	32 (10.5%)	65 (19.0%)	53 (17.4%)	106 (29.0%)
Number of migrant girls from all HHs	35	45	94	77	169
Reasons for migration (% of all migrant girls):					
<i>Boarding school</i>	14 (40%)	14 (31.1%)	23 (24.5%)	18 (23.4%)	42 (24.9%)
<i>To stay with other family member</i>	16 (45.7%)	19 (42.2%)	32 (34%)	25 (32.5%)	49 (29%)
<i>To work</i>	0 (0%)	3 (6.7%)	3 (3.2%)	5 (6.5%)	5 (3%)
<i>To get married</i>	5 (14.3%)	9 (20%)	32 (34%)	21 (27.3%)	55 (32.5%)
Migrant girl attends school	9 (25.7%)	8 (17.8%)	31 (33%)	29 (37.7%)	46 (27.2%)
Migrant girl left to another Somali village	30 (85.7%)	45 (100%)	89 (94.7%)	74 (96.1%)	146 (86.4%)
Migrant girl left Somalia	5 (14.3%)	0 (0%)	4 (4.3%)	3 (3.9%)	14 (8.3%)
<b>Sample breakdown (Boys)</b>					
Household with a migrant boy (% of total)	21 (6.1%)	25 (8.2%)	72 (21.0%)	65 (21.3%)	94 (25.8%)
Number of migrant boys from panel HH	30	30	108	101	150
Reasons for migration (% of migrant boys):					
<i>Boarding school</i>	22 (73.3%)	21 (70%)	41 (38%)	43 (42.6%)	54 (36%)
<i>To stay with other family member</i>	6 (20%)	6 (20%)	47 (43.5%)	29 (28.7%)	37 (24.7%)
<i>To work</i>	1 (3.3%)	1 (3.3%)	9 (8.3%)	12 (11.9%)	28 (18.7%)
<i>To get married</i>	1 (3.3%)	2 (6.7%)	1 (0.9%)	5 (5%)	5 (3.3%)
Migrant boy attends school	4 (13.3%)	5 (16.7%)	39 (36.1%)	32 (31.7%)	38 (25.3%)
Migrant boy left to another Somali village	28 (93.3%)	28 (93.3%)	97 (89.8%)	95 (94.1%)	124 (82.7%)
Migrant boy left Somalia	2 (6.7%)	2 (6.7%)	6 (5.6%)	2 (2%)	17 (11.3%)

## 4.2. Intersection between barriers and key characteristics

Based on the analysis above, this section identifies where the most prevalent characteristics of the sample population intersect with barriers to learning and transition. The main characteristics affecting the evaluation sample include indicators for poverty, low levels of household education, pastoralism, mental health, and the most common barriers include indicators of teaching quality and poor school governance. The tables on the

following pages present the intersection of these characteristics and barriers for cohort girls in both intervention and comparison areas enrolled in school, across evaluation rounds. It should be noted that the proportions are calculated with different denominators in each cell because the total number of respondents to each question in both rounds or areas is not the same.

Among the recorded characteristics of the cohort girls in school, in both the baseline and the midline, the most prevalent ones concern the education of primary caregivers and heads of households. Primary caregivers with no education have respectively decreased from 67.6 percent to 51.3 percent in intervention areas and 75.1 percent to 59.7 percent in comparison areas. These girls also were also faced with a number of teaching quality barriers: in the intervention areas, 44.1 percent of them also had teachers who punished students when students get lessons wrong, 41.4 percent were afraid of their teacher, 19.2 percent had teachers who treated boys and girls differently in the classroom, and 28.4 percent witnessed physical punishment last week.



Table 17: Intersection of barriers to education by subgroup characteristics for intervention areas

	Baseline	Midline	Baseline	Midline	Baseline	Midline	Baseline	Midline	Baseline	Midline	Baseline	Midline	Baseline	Midline
<b>Barriers:</b>	Home uses poor roofing material		Primary caregiver has no education		HoH has no education		Pastoralist		Female-headed household		HH dependent on charity		Mental health disability	
<b>Teachers punish when students get lessons wrong</b>	79 (25.4%)	106 (32.7%)	155 (49.8%)	143 (44.1%)	140 (47.9%)	124 (39.7%)	21 (6.8%)	15 (4.6%)	101 (32.5%)	114 (35.2%)	73 (23.5%)	81 (25%)	10 (9.8%)	33 (10.2%)
<b>Girl afraid of teacher</b>	64 (20.6%)	101 (31.2%)	116 (37.4%)	134 (41.4%)	108 (37.1%)	124 (39.7%)	16 (5.2%)	15 (4.6%)	78 (25.2%)	116 (35.8%)	61 (19.7%)	73 (22.5%)	13 (12.7%)	35 (10.8%)
<b>Teachers treat boys and girls differently in the classroom</b>	43 (14%)	47 (14.8%)	83 (26.9%)	61 (19.2%)	84 (29.1%)	58 (19%)	8 (2.6%)	6 (1.9%)	56 (18.2%)	72 (22.6%)	43 (14%)	47 (14.8%)	13 (13.1%)	25 (7.9%)
<b>Teachers often absent from class</b>	30 (9.7%)	21 (6.8%)	69 (22.3%)	17 (5.5%)	71 (24.5%)	17 (5.7%)	9 (2.9%)	4 (1.3%)	46 (14.9%)	25 (8.1%)	41 (13.3%)	21 (6.8%)	8 (7.9%)	13 (4.2%)
<b>Physical punishment witnessed last week</b>	26 (18.8%)	40 (22.7%)	56 (40.6%)	50 (28.4%)	49 (38.3%)	45 (26.6%)	4 (2.9%)	3 (1.7%)	39 (28.3%)	36 (20.5%)	31 (22.5%)	33 (18.8%)	3 (7.3%)	11 (6.3%)
<b>Communities with poor school management/poor principal</b>	5 (1.5%)	7 (2.2%)	14 (4.3%)	6 (1.9%)	7 (2.3%)	4 (1.3%)	4 (1.2%)	1 (0.3%)	5 (1.5%)	5 (1.6%)	4 (1.2%)	6 (1.9%)	0 (0%)	1 (0.3%)
<b>Communities with no active CECs</b>	51 (16.8%)	74 (24.9%)	101 (33.3%)	106 (35.7%)	101 (35.4%)	89 (30.7%)	19 (6.3%)	5 (1.7%)	73 (24.1%)	72 (24.2%)	64 (21.1%)	51 (17.2%)	11 (11.1%)	24 (8.1%)

Table 18: Intersection of barriers to education by subgroup characteristics for comparison areas

	Baseline	Midline	Baseline	Midline	Baseline	Midline	Baseline	Midline	Baseline	Midline	Baseline	Midline	Baseline	Midline
<b>Barriers:</b>	Home uses poor roofing material		Primary caregiver has no education		HoH has no education		Pastoralist		Female-headed household		HH dependent on charity		Mental health disability	
<b>Teachers punish when students get lessons wrong</b>	76 (28.5%)	97 (34.8%)	162 (60.7%)	137 (49.1%)	128 (51.8%)	110 (40.6%)	15 (5.6%)	18 (6.5%)	98 (36.7%)	106 (38%)	67 (25.1%)	79 (28.3%)	12 (13.3%)	30 (10.8%)
<b>Girl afraid of teacher</b>	55 (20.7%)	101 (36.3%)	120 (45.1%)	145 (52.2%)	95 (38.6%)	120 (44.4%)	7 (2.6%)	19 (6.8%)	71 (26.7%)	102 (36.7%)	47 (17.7%)	74 (26.6%)	10 (11.2%)	28 (10.1%)
<b>Teachers treat boys and girls differently in the classroom</b>	34 (12.8%)	49 (18.2%)	89 (33.5%)	66 (24.5%)	77 (31.2%)	58 (22.2%)	7 (2.6%)	14 (5.2%)	56 (21.1%)	62 (23%)	39 (14.7%)	48 (17.8%)	10 (11.1%)	26 (9.7%)
<b>Teachers often absent from class</b>	32 (12%)	30 (11.4%)	85 (31.8%)	38 (14.4%)	71 (28.7%)	31 (12.1%)	5 (1.9%)	13 (4.9%)	53 (19.9%)	39 (14.8%)	48 (18%)	39 (14.8%)	7 (7.8%)	15 (5.7%)
<b>Physical punishment witnessed last week</b>	25 (17.7%)	36 (21.4%)	52 (36.9%)	55 (32.7%)	36 (28.3%)	48 (28.9%)	3 (2.1%)	7 (4.2%)	31 (22%)	39 (23.2%)	19 (13.5%)	32 (19%)	2 (4.1%)	10 (6%)
<b>Communities with poor school management/poor principal</b>	10 (3.4%)	1 (0.4%)	15 (5.1%)	1 (0.4%)	8 (2.9%)	1 (0.4%)	1 (0.3%)	1 (0.4%)	5 (1.7%)	0 (0%)	8 (2.7%)	0 (0%)	1 (1%)	0 (0%)
<b>Communities with no active CECs</b>	34 (12.5%)	54 (21.4%)	77 (28.2%)	78 (31%)	65 (25.3%)	67 (27.5%)	9 (3.3%)	13 (5.2%)	56 (20.5%)	62 (24.6%)	40 (14.7%)	38 (15.1%)	6 (6.5%)	16 (6.3%)

Indicators of household poverty such as the use of poor roofing materials and dependence on charity also strongly intersect with teaching quality indicators, particularly corporal punishment by and fear of teachers. In intervention areas for the midline evaluation, 32.7 percent of girls live in houses with poor roofs and have teachers who punish students when they get lessons wrong, and 31.2 percent of girls with poor roofs also were afraid of their teachers. In comparison areas during the midline evaluation, 34.8 percent of girls were found to live in houses with poor roofing materials and have teachers who punished students when they got lessons wrong and 36.3 percent of them are afraid of teachers.

Low socioeconomic status indicators such as coming from a female-headed household also intersects with teaching quality. Girls living in female-headed household are also more frequently afraid of their teachers in the midline than in the baseline, as their proportion has increased by approximately 10 percentage points in both intervention and comparison areas.

The lack of active CECs is also a prominent barrier to education within the target communities: 24.9 percent of girls also live under poor roofs and 35.7 percent of girls also have uneducated caregivers. In comparison areas, 21.4 percent indicate that they are afraid of their teachers and 31 percent of girls have uneducated primary caregivers in addition to lacking an active CEC in their community.

The reasons for the correlation between poverty and socioeconomic indicators with poor teaching quality and poor school management are unclear, but not unexpected. Poorer communities that are unable to pay their teachers as described above are likely not able to attract trained teachers who are trained in more effective classroom management techniques that do not resort to corporal punishment or who are willing to be paid insufficient and/or infrequent salaries. With fewer trained teachers, children from poor communities may disproportionately have teachers using corporal punishment. In addition, households with limited socioeconomic status may in turn only be able to send their girls to schools with low teacher quality where punishment and poor school management may be common.

### **4.3. Appropriateness of project activities to the characteristics and barriers identified**

SOMGEP-T project activities are directed to achieve four main goals: 1) improving access to post-primary options, (2) fostering supportive school practices and conditions for marginalized girls, (3) promoting positive shifts on gender and social norms, and (4) enhancing the ability of MoEs to deliver quality education. Each of these activities address the key barriers faced by in-school, intervention cohort girls with the exception of the first which is designed to broaden educational opportunities for cohort girls who are out of school. Developing supportive school practices by teachers help marginalised girls engage with learning without fear of punishment or of the teacher. Yet, the midline evaluation shows that there have been few improvements, mainly in school facilities and family support for girls' schooling as well as adults' education. Girls still seem to fear and are still affected by poor teaching quality, including different gender treatment and frequent unnecessary and physical punishments.

Several of the prevalent barriers identified by the midline evaluation correspond with the project's ToC. These include demand-side barriers such as poverty and high chore burdens. Supply-side barriers described in the ToC and identified in the analysis include limited access to qualified teachers who in addition to teaching numeracy and literacy can increase the use of supportive, gender-sensitive learning practices and limited capacity of school leaders and education officials to address absenteeism. As poverty and land ownership have partially improved and will continue to improve throughout the program,

CARE International and its partners should continue to intervene on the supply-side barriers to address prominent issues already present in the ToC, such as teaching quality and training.

### **Project's contribution**

The majority of the findings confirm SOMGEP-T's Theory of Change and baseline assessment. The presence of minor and few improvements from the baseline evaluation provides important input to CARE International and its partners to further refine/ adjust the ToC as well as suggests that there should be more intensive targeting of the education system and teaching quality.

As suggested by the baseline report, the sample size of girls questioned regarding mental health has been expanded to include girls older than 12. The midline evaluation confirms the assumption that the whole sample of girls would have resulted in an increased proportion of girls affected by mental health issues such as depression and/or anxiety. Indeed, 11 percent of girls in both intervention and comparison areas have confirmed that they are experiencing such signs of mental health disability daily, weekly or monthly. There are still clear opportunities for addressing this point – by working with teachers to increase awareness during coaching sessions; linking with other organizations to develop simple strategies that can be adopted by teachers and girls'/ boys' empowerment fora to support students/participants who are facing anxiety and depression. It is also key to work with teachers to unpack the effects on learning.

On another note, poverty among households has improved as more families can afford to support girls' education expenses. Both girls and primary caregivers feel safer travelling to school or at school, and girls seem to endure less heavy chore burden at home. The intersection of education within households and the experience of physical discipline or discrimination in class confirms positive improvement within girls' families compared to the baseline, but the persistence of certain supply-side learning barriers concerning teachers and school facilities reaffirms the dynamics of exclusion considered in the ToC. As characteristics of discrimination and barriers of exclusion at times show positive correlation for the midline, there is a level of certainty that some dynamics identified in the ToC have not improved or been addressed properly yet, highlighting the need for work with MoEs and schools to increase teaching quality and generate better education outcomes.

## **5. Key Outcome Findings**

### **5.1. Learning Outcome**

This section presents key findings on learning outcomes for literacy and numeracy. Readers should bear in mind that the analysis in this section and all outcomes analysis below uses the full midline sample with both girls who were successfully re-contacted at midline, as well as replacement girls. First, midline targets set at the baseline are presented on the basis of the Outcomes Spreadsheet calculations. The next subsection then presents diagnostic and summary statistics as well as aggregate midline scores – including overall scores and scores by grade. The final subsections present an analysis of differences between baseline and midline as well as between intervention and comparison (i.e. difference-in-differences), as well as an analysis of learning by subtask in order to identify skill gaps. For more information on the test subtasks, marking, and scoring, please see Annex 14.

## Midline Targets

The following table presents midline targets for literacy and numeracy for the intervention group, by grade. These are targets that the girls in the intervention group are hoped to achieve over and above girls in the comparison group, thus these are the targets set for difference in differences analysis below.

Grade	Literacy Midline Target (over and above comparison group)	Numeracy Midline Target (over and above comparison group)
Grade 3 (4)	7.1	6.5
Grade 4 (5)	6.2	6.8
Grade 5 (6)	6.2	7.2
Grade 6 (7)	5.7	5.5
Overall	6.4	6.5

A summary of progress against these targets is presented in the table below:

Grade	Literacy Midline	Numeracy Midline
Target (over and above comparison group)	6.4%	6.5%
Progress against target	-20%	-12%

## Equating Baseline-Midline Assessment Difficulty

As part of the administration of midline assessments, there were 42 girls who took both the baseline and midline (Somali) literacy assessments so that their scores could be compared and the relative difficulty of the two assessments could be established. The aggregate scores of these 42 girls were as follows:

- Baseline literacy: 62.48
- Midline literacy: 63.46
- Difference: -.976 ( $p = 0.70$ )

This aggregate difference is sufficiently small as to be analytically inconsequential. However, it is also worth noting that the subtask analysis (below) revealed that a significant proportion of learners made negative progress in terms of their proficiency levels on literacy subtask 2. The tables below compare scores by subtask for both numeracy and literacy in order to identify any subtasks with significant differences in difficulty from baseline to midline:

### Numeracy Comparison

Subtask	Baseline	Midline	Difference	P-Value
1	61.6	57.4	-4.2	.50
2	93.4	90	-3.4	.32
3	82.1	85.3	3.2	.64
4	73.7	75.3	1.6	.84
5	60.5	67.9	7.4	.40
6	77	82.9	5.9	.35
7	64.2	73.2	8.9	.23
8	27.4	25.8	-1.6	.85
9	40.5	40	-0.5	.96
10	20	15.3	-4.7	.55
11	56.6	39.5	-17.1	.11

There are no numeracy subtasks with a significant difference in difficulty, so this is not a concern.

#### Literacy Comparison

Subtask	Baseline	Midline	Difference	P-Value
1	56.4	89.8	33.4	>.01
2	83.9	73.2	-10.7	.09
3	76.2	75	-1.2	.85
4	68.5	68.9	.4	.95
5	53	57.1	4.2	.60
6	51.6	54.8	3.2	.73
7	57.1	47.6	-9.5	.36
8	53.2	41.3	-11.9	.25



With regard to literacy, subtasks 1 and 2 were both significantly different in difficulty from baseline to midline. Subtask 2 was indeed more difficult at the midline than at baseline, thus accounting for the negative progress observed, while subtask 1 was easier at midline than at baseline.

These differences in subtask-1 and subtask-2 difficulty affect comparison and intervention girls equally, and thus do not bias difference-in-differences estimation. For the sake of simplicity and ease of interpretation, the analysis below proceeds without the use of an equating coefficient, and uses the standard aggregation method. The difference in subtask difficulty is borne in mind when analysing proficiency by subtask.

To be clear, the differences in difficulty observed here were not the product of intentional design and are also not immediately obvious from a direct comparison of the baseline and midline questions that contributed to these observed differences. While the reasons for these differences are unclear, the results (in terms of under-performance on the subtasks examined above) are sufficiently large that they present the possibility that students in Somalia (for reasons that are not clear) found the midline subtasks to be somewhat more difficult than the corresponding baseline subtasks.

### Aggregate Summary Statistics

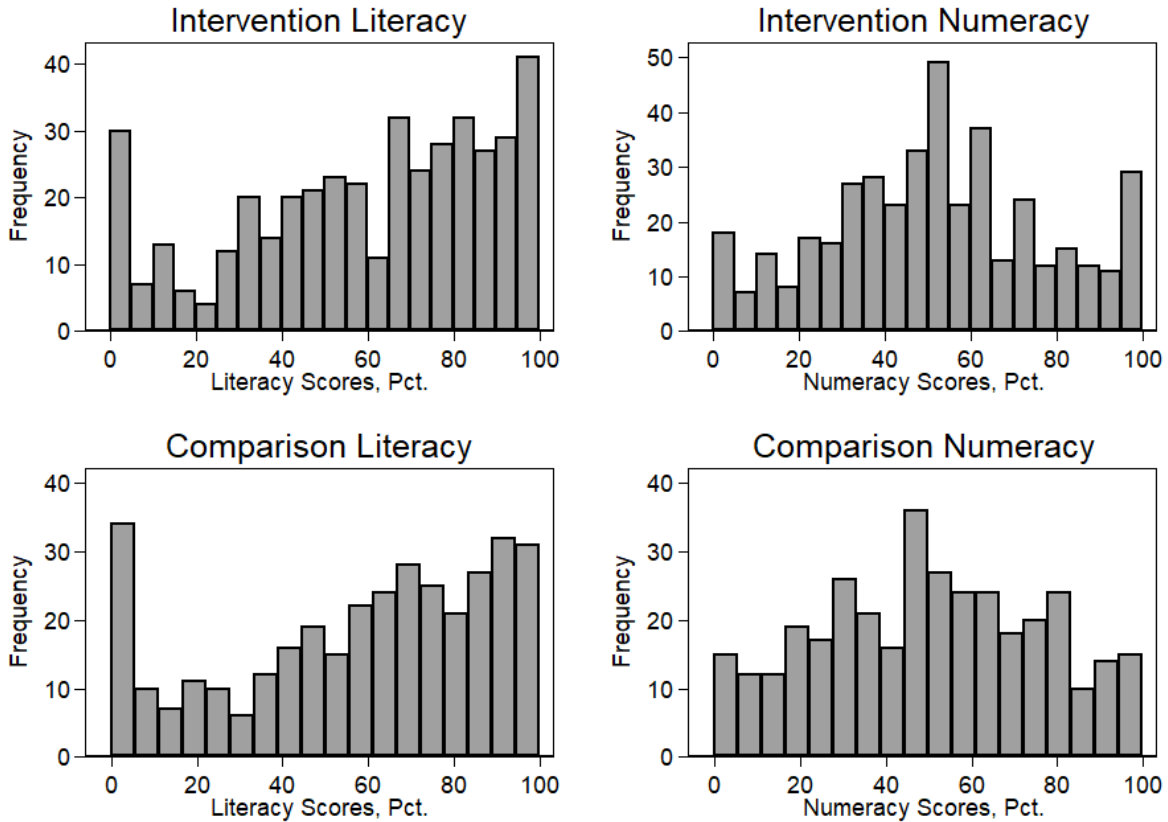
A brief summary of aggregate scores is presented here in order to validate the midline scores and explore possible floor and ceiling effects. In reviewing the distributions of baseline scores, each score was first reviewed in terms of its reliability using Cronbach’s alpha, which tests for the degree of inter-item correlations within each assessment. The results are summarized in the table below:

<i>Assessment</i>	<i>Literacy</i>	<i>Numeracy</i>
<i>Alpha</i>	0.91	0.91
<i>Internal consistency</i>	High	High

The Cronbach’s alpha is very close to 1 for each of the assessments, indicating that the level of internal consistency is high.

The panel of graphs below presents histograms of aggregate scores for literacy and numeracy for all in-school cohort girls, disaggregated by intervention and comparison groups. The total midline sample of cohort girls is N=807, with 94.9% (N=766/807) reporting that they were enrolled at the time of midline data collection. Note that the 41 girls who reported being out of school at the time of the midline are not included in the histograms below.

Figure 2: Distribution of learning scores by intervention versus comparison



Literacy and Numeracy scores do not show significant evidence of significant floor or ceiling effects. Baseline literacy scores had exhibited extreme floor effects in English literacy, which contributed to overall floor effects in the aggregate literacy scores at the baseline. The removal of English literacy from the midline assessments has resulted in literacy scores that have a much less problematic distribution. On the low end, fewer than 5% of girls scored lower than 2% on the literacy portion of the assessment. On the high end, fewer than 4% of girls scored between 98% and 100%. Literacy scores for the Intervention group exhibit a slight degree of bimodality (with scores showing a modest level of clustering at the top and bottom of the 100-point range). This slight bimodality should not create problems for difference-in-differences analysis, but will nonetheless be borne in mind interpreting results below. Numeracy scores are well distributed and raise no concerns related to floor or ceiling effects, or bimodality.

The analysis of learning outcomes below focuses on the scores of in-school girls in grades 3-6 (3-7), since this was the targeted and expected grade-range for midline evaluation. The inclusion of grade 3 (3) girls allows for the fact that some girls who belonged to grade 3 at the baseline were properly re-sampled as part of the midline panel, but had not advanced a grade since the baseline and thus remained in grade 3 (3). As noted earlier, the sample also inadvertently included 41 girls who reported being out of school at the time of the midline, as well as 12 girls whose grade-levels were reported as below grade 3 (which should not be possible) and whose grade levels could not be established with certainty despite follow-up

attempts with caretakers and school administrators. Finally, there were 8 girls who reported being in grade 7 (8) at the midline. Each of these categories of girls – OOS, grade level uncertain, and grade 7 (8) – are considered non-comparable from baseline to midline, and thus are excluded from the analysis below.

## Literacy

The aggregate Somali literacy scores for the midline sample are 59.8 for the intervention group and 58.5 for the comparison group. The table below presents literacy results by grade and by intervention versus comparison groups (with aggregates in the final row). The table shows that learning outcomes and grades are correlated, with increased grade-level predicting increased learning. Mean literacy scores increase monotonically (i.e. stepwise) by grade within the intervention group. The comparison group shows monotonic increases in the main cohort grades 3-6 (4-7), but girls in grades 2 (3) and 3 (4) do not fit the expected pattern, with the average scores for grade 3 (3) being higher than the average scores for girls in grade 3 (4). It is likely that these uncharacteristically high scores for girls in grade 3 (3) in the comparison group are merely a result of the very small sizes of these subsamples (with 26 grade 3 (3) girls in the comparison group).

It is worth noting that, with the exception of grade 3 (3) girls, the mean literacy scores for the intervention group are consistently higher than the mean literacy scores for the comparison group. This finding provides preliminary evidence that girls in the intervention group may be out-performing girls in the comparison group as a result of the project’s interventions. A more robust analysis of this hypothesised effect of project interventions will be presented in the longitudinal comparisons made below.

Finally, as we look at increases in learning across grades, there is a leap in intervention-group literacy scores from grade 4 (5) to grade 5 (6). Average literacy increases by 12.7 percentage points from grades 4-5 (5-6) in the intervention group, while it only increases by 9.4 percentage points for the comparison group from grades 4-5 (5-6). This leap may indicate that there are certain literacy skills that intervention girls are learning from grade 4 (5) to 5 (6) that their peers in the comparison group are not learning as readily. This finding is also consistent with the focus, in this phase of the project, on higher level literacy skills. This question of differential skill acquisition will be taken up in greater detail in the analysis of skill gaps below.

*Table 19: Literacy (EGRA/SeGRA)*

Grade	Intervention Group Mean	Comparison Group Mean	Standard Deviation in the intervention group
Grade 3 (3)	13.2	46.8	19.4
Grade 3 (4)	51.5	43.6	28.4
Grade 4 (5)	58.7	58.2	25.7
Grade 5 (6)	71.4	67.5	20.7
Grade 6 (7)	77.3	72.2	19.5
Overall	59.8	58.5	28.6

The table below shows changes in literacy scores from baseline to midline and for intervention versus comparison groups, by grade. The final column of the table presents average difference in differences by grade (i.e. average change in score over time for the intervention group minus the average change in score over time for the comparison group). When disaggregated by grade, it is clear (for both the intervention and comparison groups) that, on average, girls in grades 3 to 6 (4 to 7) improved their literacy scores over time. However, the average amount of improvement shown by girls in the intervention group tended to be smaller than the average amount of improvement shown by girls in the comparison group, which is reflected in the fact that most of the average difference in differences are negative in their sign and are very close to zero. At the descriptive level this is preliminary evidence suggesting that at this preliminary stage (after four months of exposure) the intervention had no measurable positive effect on literacy at most grade levels (and in the aggregate). Grade 3 (4) is the only subgroup where the difference in differences is positive, but this change is not statistically significant. At the aggregate level, the difference in differences for literacy scores is -1.7 which can be interpreted as no intervention-effect.

The difference in differences figures are particularly large and negative for girls in grade 3 (3), which suggests that there has been significant stagnation and even possibly some negative progress or reductions in performance in terms of literacy learning for girls in grade 3 (3). For the 52 girls in grade 3 (3) at the midline, these results can be attributed to the fact that nearly all of these girls were either asked to repeat grade 3 (meaning they were in grade 3 at baseline and midline), or were demoted a grade (meaning that they were in grade 4 at baseline and are now in grade 3 at the midline).

Grade	Baseline Literacy Intervention	Midline Literacy Intervention	Difference Baseline to Midline (Intervention)	Baseline Literacy Comparison	Midline Literacy Comparison	Difference Baseline to Midline (Comparison)	Difference in Difference (Intervention Diff – Comparison Diff)
Grade 3 (3)	12.2	13.2	1.0	29.5	46.8	17.2	-16.2
Grade 3 (4)	30.5	51.5	20.9	24.8	43.6	18.8	2.1
Grade 4 (5)	43.7	58.7	15.1	39.8	58.2	18.4	-3.3
Grade 5 (6)	57.1	71.4	14.3	52.3	67.5	15.2	-0.9
Grade 6 (7)	58.2	77.3	19.2	49.9	72.2	22.2	-3.1
Overall	43.4	59.8	16.4	40.4	58.5	18.1	-1.7

For reference, the tables below present composition of the sample in terms of grade-level and the status of being held back or repeating a grade, as well as by intervention versus comparison groups. The first table makes it clear that grade 3 (3) girls comprise the highest number and proportion of demoted and held-back girls in the sample. The second table shows that the proportion of held-back and demoted girls in the sample does not differ significantly by intervention versus comparison groups.

Grade	Number Successful Transition	Number Demoted	Number Held Back
Grade 3 (3)	0	11	41

<b>Grade 3 (4)</b>	178	6	20
<b>Grade 4 (5)</b>	146	4	26
<b>Grade 5 (6)</b>	165	0	19
<b>Grade 6 (7)</b>	130	0	0
<b>Total</b>	619	21	106

<b>Grade</b>	<b>Number Successful Transition (Row %)</b>	<b>Number Demoted (Row %)</b>	<b>Number Held Back (Row %)</b>
<b>Comparison</b>	282 (83.2)	8 (2.4)	49 (14.5)
<b>Intervention</b>	337 (82.8)	13 (3.2)	57 (14.0)
<b>Total</b>	619 (83.0)	21 (2.8)	106 (14.2)

These girls who were held back or demoted a grade were, in many cases, already performing below average for their grade level at the baseline, and it is likely that many were held back or asked to repeat a grade for that very reason. It is worth noting that the qualitative data provides another important reason why girls may have been asked to repeat a grade, namely armed conflict resulting in extended school closures. As one teacher explained the problem: "...this village was in conflict and because of this a full academic year was lost. For six months, students were not able to take exams, but since eight months ago all students have been called to attend school again. The students we have now are at the same grade as last year but the situation has now improved."<sup>63</sup> It is important to note that the teacher's explanation links armed conflict not only to the significant loss of instructional time, but also to girls repeating the same grade as the previous year, thus providing a clear linkage among conflict, attenuated learning, and girls repeating a grade. In the subgroup and barriers analysis below, special attention will be given to the population of held-back and demoted girls, examining the degree to which demographics and associated barriers are associated with girls falling into this high-risk category.

Over all, the descriptive analysis thus far suggests that on average girls tended to improve their literacy levels over time as a result of being in school, but that the intervention has not had a significant positive effect on literacy learning over and above the effect that would be expected from simply going to school in the absence of the intervention. The literacy regression results presented in the table below are a direct statistical test of the hypothesis that the intervention contributed to an increase in literacy outcomes. This statistical test leaves out OOS girls because OOS girls in the intervention group dropped out during the first year of the intervention, and thus are not expected to exhibit an intervention-effect over and above OOS girls in the comparison group. It should also be noted that, whereas all results presented above include replacement girls, the regression results below exclude replacements, and are thus estimated only on the basis of the 'true' panel sample of girls who were successfully re-contacted at midline.

---

<sup>63</sup> FGD – Teachers

Table 3b: Literacy results<sup>64</sup>

Result	Details	Comments
Literacy Baseline - Midline	Beta = -1.29 p-value <sup>65</sup> = 0.70 Target = 6.4 Performance against target <sup>66</sup> = -20% N = 564 panel girls	Regression is limited to cohort girls in grades (3-7), including girls who repeated grade 3 at baseline. The estimated intervention effect (difference-in-differences coefficient) is negative and not statistically significant. The intervention has not had a measurable positive effect on literacy learning, and intervention girls have made negative progress toward the midline target.

The difference-in-differences results presented in the table above confirm that there has been no detectable positive effect of the intervention on literacy learning. As in the summary tables above, the regression results also suggest that the amount of improvement since baseline in the comparison group has been larger than the amount of improvement in the intervention group, giving a negative coefficient in the regression. Ultimately, the net effect of the intervention since the baseline is indistinguishable from zero, and girls in the intervention group have made negative progress against the target of a 6.4 percentage point improvement (over and above the comparison group) since the baseline.

### Numeracy

The aggregate numeracy scores for the midline sample are 51.9 for the intervention group and 50.6 for the comparison group. The table below presents numeracy results by grade (including out-of-school girls) and by intervention versus comparison groups (with aggregates in the final row). As with literacy scores, numeracy learning outcomes and grades are correlated, with increased grade-level predicting increased learning.

As with literacy scores, there is a major leap in performance by, 10.4 percentage points, from grade 4 (5) to grade 5 (6) girls in the intervention group, as compared with an increase of 7.7 percentage points from grade 4 (5) to grade 5 (6) in the comparison group. This leap may indicate that there are certain numeracy skills that intervention girls are learning from grade 4 (5) to 5 (6) that their peers in the comparison group are not learning as readily. This question of differential skill acquisition will be taken up in greater detail in the analysis of skill gaps below.

Table 20: Numeracy (EGMA/SeGMA)

Grade	Intervention Group Mean	Comparison Group Mean	Standard Deviation in the intervention group

<sup>64</sup> Results reported here are for 595 girls in the baseline who correspond to 595 successfully re-contacted girls in the midline. This estimation sample thus excludes replacement girls. The substantive results of the regression do not change when the estimation sample is expanded to include replacement girls. Please see Annex 3 for a detailed summary of all regression results.

<sup>65</sup> Beta and p-values are estimated using an unweighted, linear regression with cluster-robust standard errors.

<sup>66</sup> This represents adjusted performance against target calculated in the Outcomes Spreadsheet on the basis of the adjusted difference-in-differences estimate.



<b>Grade 3 (3)</b>	25.0	40.5	21.8
<b>Grade 3 (4)</b>	45.1	41.4	23.4
<b>Grade 4 (5)</b>	51.1	47.8	22.1
<b>Grade 5 (6)</b>	61.5	55.5	24.0
<b>Grade 6 (7)</b>	61.8	65.1	23.1
<b>Overall</b>	51.9	50.6	25.0

The table below shows changes in numeracy scores from baseline to midline and for intervention versus comparison groups, by grade. The final column of the table presents average difference in differences by grade. When disaggregated by grade, it is clear (for both the intervention and comparison groups) that, on average, girls in all grade levels improved their numeracy scores over time. However, the average amount of improvement shown by girls in the intervention group tended to be smaller than the average amount of improvement shown by girls in the comparison group, which is reflected in the fact that most of the average difference in differences by grade are negative in their sign and are very close to zero. At the descriptive level this is preliminary evidence suggesting that at this initial stage of exposure to the intervention (four months), the intervention had no measurable positive effect on literacy at most grade levels (and in the aggregate). Grades 3 (4) and 5 (6) are the only subgroups where the difference in differences are positive (albeit not statistically significant, and the intervention group girls in grade 5 (6) have met and exceeded their numeracy target of 7.2 by 0.6 percentage points. At the aggregate level, the difference in differences for numeracy scores is -0.6 which can be clearly interpreted as no positive, measurable intervention-effect.

<b>Cohort</b>	<b>Baseline Literacy Intervention</b>	<b>Midline Literacy Intervention</b>	<b>Difference Baseline to Midline (Intervention)</b>	<b>Baseline Literacy Comparison</b>	<b>Midline Literacy Comparison</b>	<b>Difference Baseline to Midline (Comparison)</b>	<b>Difference in Difference (Intervention Diff – Comparison Diff)</b>
<b>Grade 3 (3)</b>	20.6	25.0	4.4	30.9	40.5	9.5	-5.2
<b>Grade 3 (4)</b>	32.0	45.1	13.2	28.4	41.4	13.0	0.2
<b>Grade 4 (5)</b>	39.2	51.1	11.9	33.5	47.8	14.3	-2.4
<b>Grade 5 (6)</b>	46.1	61.5	15.4	47.9	55.5	7.7	7.8
<b>Grade 6 (7)</b>	53.4	61.8	8.4	45.9	65.1	19.2	-10.8
<b>Overall</b>	40.0	51.9	11.9	38.1	50.6	12.5	-0.6

As with literacy scores above, difference in differences figures are particularly large and negative for girls in grade 3 (3), which suggests that there has been stagnation or potentially even negative progress or reductions in performance in terms of numeracy learning for girls in this subgroup. As above, these results can be attributed to the fact that all of these girls were either asked to repeat grade 3 or were demoted a grade, which itself indicates that their teachers probably realized that those girls were learning much slower than their peers. It is likely that these girls were already in the process of falling behind as of the baseline because they were already below grade-level in their performance at baseline, and the

midline results provide even clearer evidence that these girls are at risk because they are losing ground year-on-year, and are probably also at risk of dropping out of school altogether. The causes of this negative progress will be explored in greater detail in the analysis of subgroups and barriers below.

Result	Details	Comments
<b>Numeracy Baseline - Midline</b>	Beta = -0.80 p-value <sup>67</sup> = 0.85 Target = 6.5 Performance against target <sup>68</sup> = -12% N=564 panel girls	Regression is limited to cohort girls in grades 3-6 (3-7). The estimated intervention effect (difference-in-differences coefficient) is negative and not statistically significant. The intervention has not had a measurable positive effect on numeracy learning, and intervention girls have made negative progress toward the midline target.

The difference-in-differences results presented in the table above confirm that there has been no detectable positive effect of the intervention on numeracy learning. As in the summary tables above, the regression results also suggest that the amount of improvement since baseline in the comparison group has been approximately the same as the amount of improvement in the intervention group, giving a regression coefficient that is negative and very close to zero. Ultimately, the net effect of the intervention since the baseline is indistinguishable from zero, and girls in the intervention group have made negative progress against the target of a 6.5 percentage point improvement (over and above the comparison group) since the baseline.

### Identifying Foundational Skill Gaps

This section identifies potential skill gaps through an analysis of learning outcomes by subtask and by achievement category (non-learner, emergent learner, established learner, and proficient learner). The tables below present the percentage of in-school, cohort girls in the intervention group (n=419) who fall into a given learning category for a given subtask.<sup>69</sup> Note that the sample of learners presented here includes five girls whose midline grades were unverifiable (and who were thus omitted from the grade-wise analysis above).

The tables on the following page present foundational skill gaps for numeracy. As expected, the percentage of non-learners generally increases as a function of increasing sub-task difficulty, while the percentage of proficient learners decreases correspondingly. At higher levels of difficulty, the distribution

<sup>67</sup> Beta and p-values are estimated using an unweighted, linear regression with cluster-robust standard errors.

<sup>68</sup> This represents adjusted performance against target calculated in the Outcomes Spreadsheet on the basis of the adjusted difference-in-differences estimate.

<sup>69</sup> All tabulated results are weighted to adjust for the fact that the number of girls in each cluster varies widely (from 4 to 21). Weights are applied such that each school-cluster counts evenly toward the estimated percentages. Thus, the school-level weight is  $\frac{21}{n}$  where  $n$  = the number of in-school girls in a given school-cluster. This weighting mimics the weighting used at baseline and is necessary in order to avoid a possible scenario in which the number of girls in a given cluster is correlated with learning outcomes.

of achievement levels shows less evidence of bimodality than at the baseline, where the number of emergent and established learners became smaller and smaller at higher skill levels.

At the baseline, the primary skill gap in numeracy emerged between simple addition and subtraction (subtask 2 to subtask 3). This skill gap has completely disappeared at the midline, with the number of proficient learners in subtraction having increased by 25.0 percentage points since the baseline to the point that the proportion of proficient learners in basic subtraction is effectively equivalent to the proportion of proficient learners in basic addition. Correspondingly, the proportion of non-learners in subtasks 1-3 is now consistently below 10%, indicating that only a small minority of learners remain stuck in a place where they lack the fundamental understanding of arithmetic that will allow them to continue to learn higher-level numeracy skills.

The primary skill gap at midline is now between basic subtraction (subtask 3) and more advanced addition (subtask 4). This skill gap is already being closed gradually, with the proportion of proficient learners having increased by 13.0 percentage points in subtask 4 since the baseline. It is not clear whether remedial work is necessary in order to address this skill gap, but it appears that there has been some improvement since baseline, merely as a result of students improving more fundamental skills at the level of subtasks 2 and 3.

Table 21: Foundational numeracy skills gaps

Categories	Subtask 1	Subtask 2	Subtask 3	Subtask 4	Subtask 5	Subtask 6	Subtask 7	Subtask 8	Subtask 9	Subtask 10	Subtask 11
	Missing Number	Addition (Level 1)	Subtraction (Level 1)	Addition (Level 2)	Subtraction (Level 2)	Word problems (add/subtract)	Multiplication (Level 1)	Multiplication (Level 2)	Division (Level 1)	Division (Level 2)	Word problems (mult/div)
<b>Non-learner 0%</b>	2.3 (-1.1)	2.7 (-3.1)	6 (-18.5)	15.3 (-18.9)	23.6 (-24.4)	11.5 (-21.9)	18.1 (-28.3)	69.6 (-6.9)	49.1 (-22)	75.9 (-9.2)	58.9 (-11)
<b>Emergent learner 1%-40%</b>	37.1 (7.1)	2.8 (-1.4)	3.9 (-3.5)	15.4 (-6.1)	10.6 (-3)	5.2 (-1.9)	15.3 (0.8)	12.4 (3.9)	15 (4)	11.8 (5.6)	0 (0)
<b>Established learner 41%-80%</b>	37.8 (9.4)	17.8 (-1)	12.1 (-3)	34.1 (12)	33.9 (12.5)	33.5 (2.5)	41.8 (23)	9.5 (4.7)	16.2 (9.8)	3.9 (0.8)	14 (2.3)
<b>Proficient learner 81%-100%</b>	22.7 (-15.4)	76.7 (5.4)	78 (25)	35.2 (13)	31.9 (15)	49.8 (21.3)	24.7 (4.5)	8.4 (-1.7)	19.8 (8.1)	8.5 (2.8)	27.2 (8.7)
	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Table 22: Foundational numeracy skills gaps with mean scores by grade

Categories	Subtask 1 Missing Number	Subtask 2 Addition (Level 1)	Subtask 3 Subtraction (Level 1)	Subtask 4 Addition (Level 2)	Subtask 5 Subtraction (Level 2)	Subtask 6 Word problems (add/subtract)	Subtask 7 Multiplication (Level 1)	Subtask 8 Multiplication (Level 2)	Subtask 9 Division (Level 1)	Subtask 10 Division (Level 2)	Subtask 11 Word problems (mult/div)
Grade 3 (3)	32.8	60.9	59.2	18.8	22.2	38.5	24.0	0.0	12.2	0.9	15.3
Grade 3 (4)	54.1	85.8	85.8	58.3	53.7	68.1	50.1	9.8	28.2	8.9	24.8
Grade 4 (5)	55.2	90.4	88.0	66.5	61.8	75.9	60.5	17.0	34.1	15.1	29.9
Grade 5 (6)	62.3	92.1	90.6	73.3	70.7	78.2	67.5	25.8	41.0	19.3	39.8
Grade 6 (7)	66.6	92.3	90.9	76.5	67.4	85.7	71.8	32.4	47.8	21.3	55.0

Table 23: Aggregate DID analysis of mean numeracy scores by subtask

Categories	Subtask 1 Missing Number	Subtask 2 Addition (Level 1)	Subtask 3 Subtraction (Level 1)	Subtask 4 Addition (Level 2)	Subtask 5 Subtraction (Level 2)	Subtask 6 Word problems (add/subtract)	Subtask 7 Multiplication (Level 1)	Subtask 8 Multiplication (Level 2)	Subtask 9 Division (Level 1)	Subtask 10 Division (Level 2)	Subtask 11 Word problems (mult/div)
Intervention Baseline	66.5	81.7	62.2	43.5	35.3	49.1	34.7	15.9	17.7	9.5	22.6
Intervention Midline 1	56.2	86.5	84.6	62.4	57.1	72.2	57.2	16.6	32.6	13.2	32.3
Comparison Baseline	59.3	76.9	64.5	42.6	35.1	46.3	34.6	13.2	16.9	9.4	21.3
Comparison Midline 1	54.0	85.1	81.7	58.3	54.0	69.2	55.2	16.2	35.1	12.7	35.5
DiD	-5.1	-3.4	5.2	3.2	2.9	0.2	1.9	-2.2	-3.3	0.4	-4.5

At the baseline, a second major inflection point occurred between multiplication skill levels (subtask 7 to subtask 8), and this gap persists at the midline, with very little change in the distribution of learners in terms of their proficiency in more advanced multiplication at subtask 8. The persistence of this skill gap is increasingly problematic as girls are expected to have learned these skills by the time they reach grade 4.

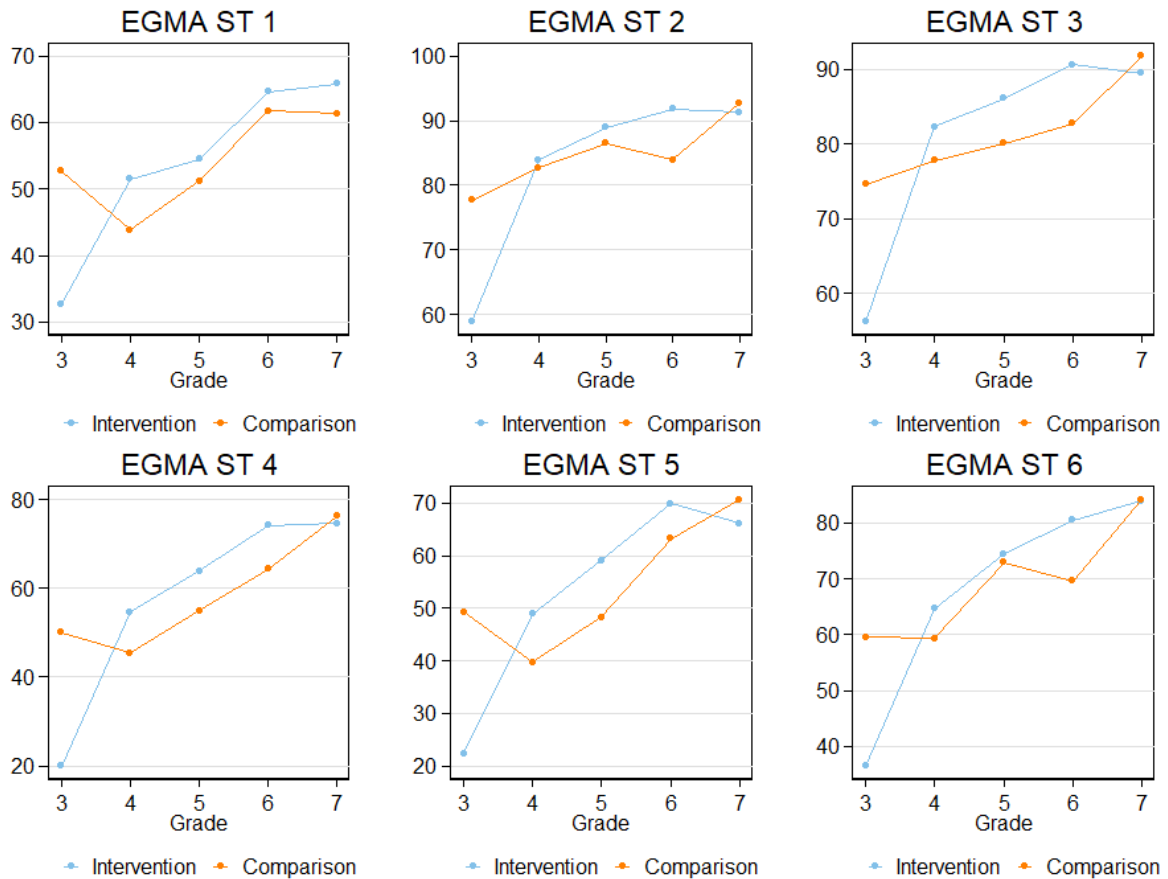
As at baseline, the achievement levels on word-problem subtasks (and the longitudinal improvements in those achievement levels) generally parallel achievement levels for the basic operations (e.g. addition or multiplication) that are most relevant to solving word problems at a given level. For example, the proportion of proficient learners in subtraction has increased by 25.0 percentage points, and correspondingly, the proportion of proficient learners in word-problems that use subtraction has increased by 21.3 percentage points, indicating that improvements in skill level with the operation (subtraction in this case) are tracking improvements in the application of that skill within a word problem. This finding implies that performance on word problems is keeping pace with overall skill development, and so this is not an area of concern.

Finally, the difference-in-differences subtask table above provides some additional insight into the overall finding that mean numeracy scores increased more in the comparison group than in the intervention group. With the exception of subtask 1 (which is exceptional because this subtask was more difficult at midline than at baseline) the trend is that the largest negative difference in differences appear where the comparison group mean was well below the intervention mean at baseline. Thus, the negative difference in differences are not a product of the comparison girls scoring higher than intervention girls at the midline; rather, this finding is a result of comparison girls who were behind at the baseline effectively catching up or closing the gap with intervention girls, while not passing them by.

In order to further explore numeracy skill gaps and to better understand the problem of plateauing performance and negative progress among girls in grade 3 (3), the panel of graphs below present numeracy scores by subtask and by grade for all in-school girls.



Figure 3: Numeracy subtasks 1-6, by grade

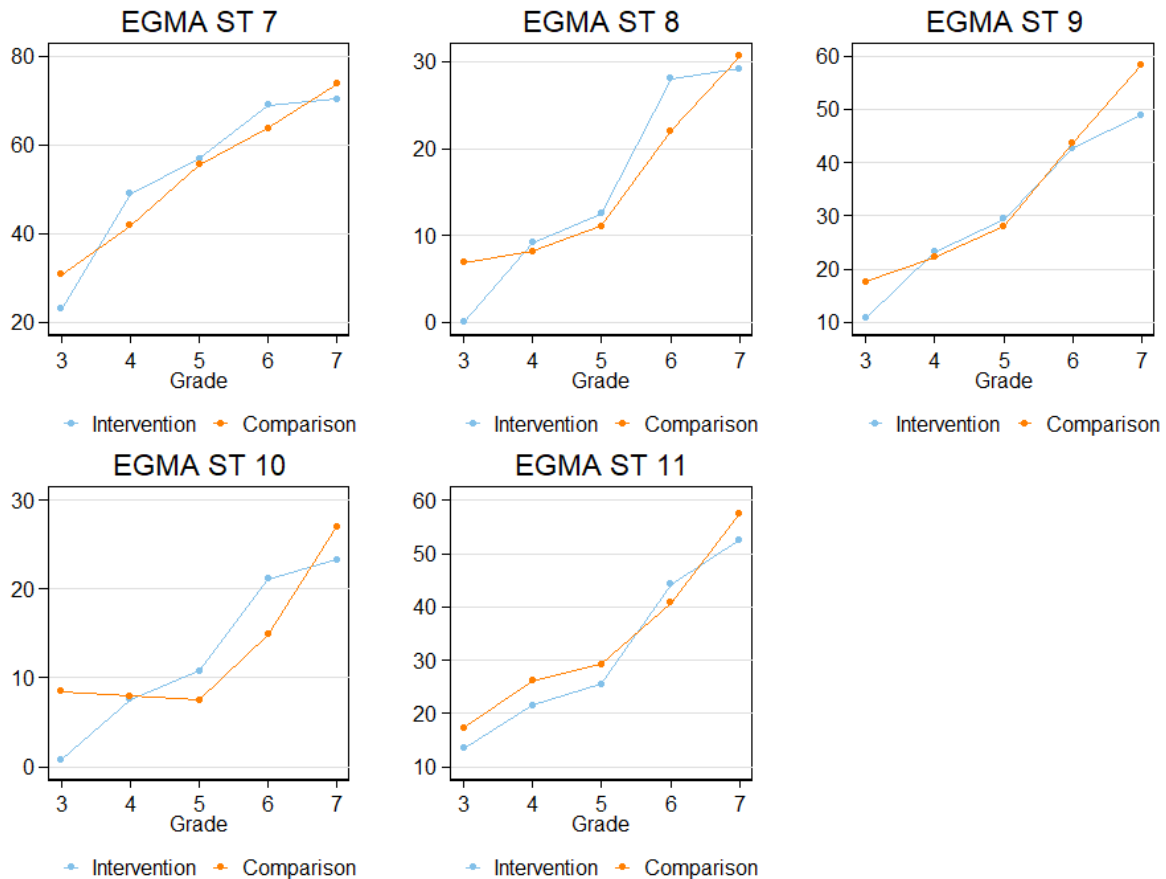


The skill gap between grade 3 (3) girls in the intervention group as compared with girls in the comparison group is immediate and extreme – meaning that the gap emerges at the most fundamental levels (subtasks 1-3) where it is in excess of 30 percentage points difference between intervention and comparison, and then persists through higher skill levels. The grade 3 (3) performance gap does eventually close at the highest skill levels where the mean scores of grade 3 (3) girls in the comparison group come down to meet the scores of grade 3 (3) girls in the comparison group. The closing of this gap is most evident from subtask 6 (on the panel of graphs above) to subtask 7 (on the panel of graphs below). This finding implies that grade 3 (3) girls in the intervention group have fallen behind in the areas of addition and subtraction, vis-à-vis their peers in the comparison group. At higher skill levels, all grade 3 (3) girls fall into the multiplication skill gap (at subtask 7), which explains why their scores converge at that point.

The grade 3 (3) skill gap here is a result of two peculiarities, with the first being the severe under-performance of grade 3 (3) girls in the intervention group, and the second being the unexpectedly high performance of grade 3 (3) girls in the comparison group. In the graphs above and below, it is clear that grade 3 (3) girls in the comparison group are performing nearly as well, or in some cases better than, grade 3 (4) girls in the comparison group. Making reference back to the baseline-midline comparison tables above, it is clear that the comparison group girls in grade 3 (3) at the midline were already out-performing their peers in terms of aggregate numeracy scores at the baseline. Since these comparison

group girls were clearly performing well as of the baseline, it is difficult to understand why they would have been held-back or demoted.

Figure 4: Numeracy subtasks 7-11, by grade



Shifting from numeracy to literacy, the tables on the following page present foundational skill gaps for Somali Literacy. By far, the largest improvement from baseline to midline is in the proportion of proficient learners in reading words (subtask 1), with an increase of 75.1 percentage points from the baseline. This increase in foundational reading skill is also reflected in the increase in reading fluency (subtask 4, passage reading) where the proportion of proficient learners has increased by 22.9 percentage points from baseline to midline.

Table 24: Foundational literacy skills gaps

Categories	Somali ST1 Reading Words	Somali ST2 Reading Comp (easy)	Somali ST3 Reading Comp (medium)	Somali ST4 Reading Fluency	Somali ST5 Reading Comp (difficult)	Somali ST6 Writing (fill blank)	Somali ST7 Writing (negative form)	Somali ST8 Writing (future tense)
<b>Non-learner 0%</b>	2.8 (-11.1)	9.8 (-5.9)	13.6 (-10.2)	4.9 (-20.4)	25.5 (-12.6)	28.2 (-17.7)	39.8 (-22.3)	46.2 (-21.5)
<b>Emergent learner 1%-40%</b>	5.2 (-41.1)	6.6 (4.4)	10.8 (3.8)	18 (-5.8)	17.2 (2.7)	17.6 (4)	5.2 (0)	7 (3.7)
<b>Established learner 41%-80%</b>	15.8 (-22.9)	40.1 (12.5)	39 (10.3)	36.6 (3.3)	42.7 (14.8)	18.3 (0.2)	7.9 (-3.3)	6.3 (-1.7)
<b>Proficient learner 81%-100%</b>	76.2 (75.1)	43.4 (-11)	36.6 (-3.9)	40.5 (22.9)	14.6 (-4.9)	36 (13.5)	47.1 (25.6)	40.6 (19.5)
	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Table 25: Foundational literacy skills gaps with mean scores by grade

Categories	Somali ST1 Reading Words	Somali ST2 Reading Comp (easy)	Somali ST3 Reading Comp (medium)	Somali ST4 Reading Fluency	Somali ST5 Reading Comp (difficult)	Somali ST6 Writing (fill blank)	Somali ST7 Writing (negative form)	Somali ST8 Writing (future tense)
<b>Grade 3 (3)</b>	37.6	14.2	8.4	14.6	8.6	12.0	8.4	3.7
<b>Grade 3 (4)</b>	81.4	65.3	58.9	54.7	44.7	46.3	44.7	37.4
<b>Grade 4 (5)</b>	87.9	77.1	67.7	65.9	43.1	51.2	50.2	41.5
<b>Grade 5 (6)</b>	92.5	82.5	72.2	76.9	47.7	66.9	61.6	55.0
<b>Grade 6 (7)</b>	94.6	83.2	79.9	85.6	58.9	69.8	81.6	76.0

Table 26: Aggregate DID analysis of mean literacy scores by subtask

Categories	Somali ST1 Reading Words	Somali ST2 Reading Comp (easy)	Somali ST3 Reading Comp (medium)	Somali ST4 Reading Fluency	Somali ST5 Reading Comp (difficult)	Somali ST6 Writing (fill blank)	Somali ST7 Writing (negative form)	Somali ST8 Writing (future tense)
<b>Intervention Baseline</b>	33.0	72.4	58.2	43.4	39.3	39.2	30.8	26.8
<b>Intervention Midline 1</b>	84.3	70.4	63.0	64.0	44.7	52.7	53.7	46.0
<b>Comparison Baseline</b>	30.5	67.0	52.5	38.7	37.1	39.4	31.6	27.8
<b>Comparison Midline 1</b>	81.3	69.2	58.3	61.9	46.5	53.9	51.8	45.4
<b>DID</b>	0.4	-4.2	-0.9	-2.7	-4.0	-1.1	2.8	1.6

Notwithstanding these improvements in foundational skills related to reading words and passages, reading comprehension (measured in subtasks 2, 3, and 5) has not improved substantially since the baseline. There has been a modest reduction in the proportion of non-learners, but there has been no corresponding increase in the proportion of proficient learners across the subtasks related to reading comprehension.

The findings for subtask 2 are, in part, driven by the fact that this subtask was more difficult at the midline than at the baseline, helping to explain why the proportion of proficient learners in subtask 2 was reduced by 11 percentage points from the baseline to the midline.<sup>70</sup> Since the findings for subtask 2 are questionable, the findings from subtask 3 can be taken as a more reliable indicator of change over time in levels of reading comprehension. Based on this comparison, it is clear that basic reading comprehension is still the area that permits the greatest room for improvement.

The difference in differences for subtasks 2 and 3 (which are not sensitive to variations in difficulty from baseline to midline) are negative, suggesting that intervention girls have not improved over and above comparison girls since the time of the baseline. The fact that the proportion of proficient learners in these subtasks has not improved substantially since the baseline is at least partly attributable to the short duration of the project intervention thus far, but it also suggests that additional remedial efforts may be necessary in order to ensure that girls whose learning is stagnating or falling behind can catch up with their peers.

At baseline, the most extreme skill gap was between medium and difficult reading comprehension (subtasks 3 and 5), and this skill gap also persists at the midline. While this persistent skill gap is worth noting, the primary skill gap remains at easier levels of reading comprehension. It is likely that if the more

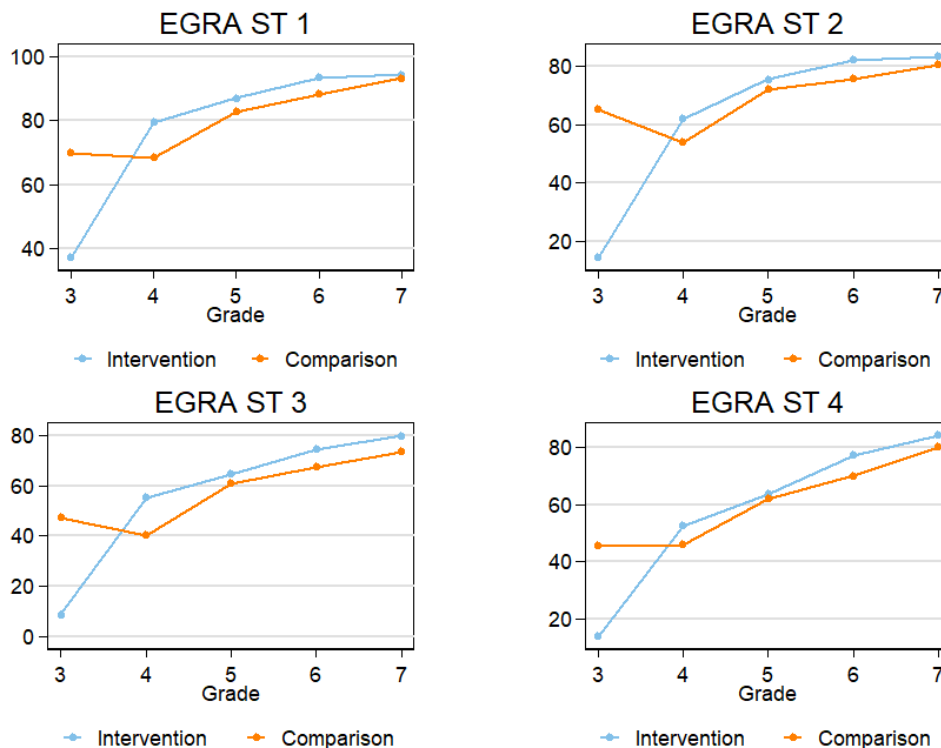
<sup>70</sup> For more information about how the relative difficulty of these tasks was established, readers can refer to the above analysis in the subsection on equating baseline-midline assessment difficulty.

fundamental skill gap is addressed, this secondary skill gap will naturally dissipate as girls have more practice.

Writing skills are generally keeping pace with reading comprehension skills, as the proportion of proficient learners in easy and medium reading comprehension is approximately the same as the proportion of proficient learners in the writing subtasks.

In order to further explore literacy skill gaps and to better understand the problem of plateauing performance and negative progress among intervention girls in grade 3 (3), the panel of graphs below present literacy scores by subtask and by grade for all in-school girls.

Figure 5: Literacy subtasks 1-4, by grade

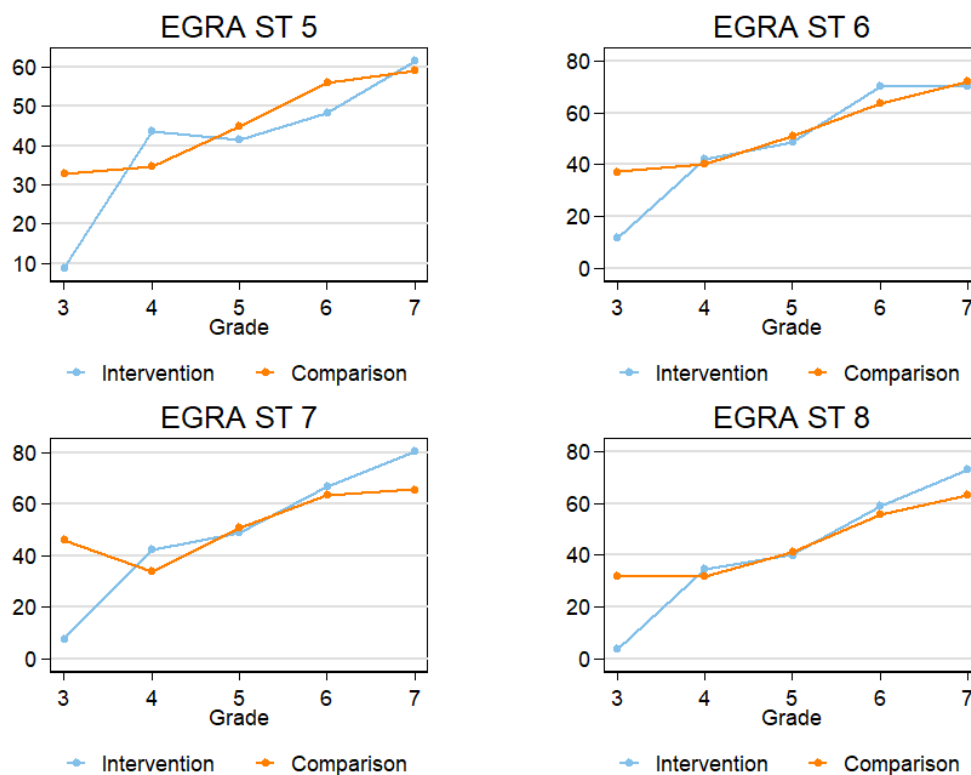


As with numeracy, the skill gap between intervention and comparison girls in grade 3 (3) is immediate and severe. At the most fundamental subtasks, the gap is in excess of 30 percentage points. Unlike the skill gap for numeracy, the literacy skill gap persists even at the highest levels of difficulty with subtasks 7 and 8, where grade 3 (3) girls in the intervention group continue to perform below what might be expected, while grade 3 (3) girls in the comparison group perform nearly as well as or in some cases better than their peers in grade 3 (4). These findings with regard to literacy provide further evidence that the grade 3 (3) skill gap is a product both of under-performance in the intervention group and higher-than-expected performance in the comparison group.

With regard to the leap in intervention-group learning from grade 4 (5) to grade 5 (6), the greatest subtask-level increases are in the skills of reading fluency and writing (subtasks 4, and 6-8), which are

higher level literacy skills that are in the process of being addressed through project interventions at the school level.

Figure 6: Literacy subtasks 5-8, by grade



### Boys' data and comparison with girls' data

Analysis of boys' learning outcomes reveals significant gender disparities, although the aggregate learning gap between boys and girls has been reduced since the baseline. In particular, girls in grade 3 (4) are abruptly closing the learning gap in numeracy and have already overtaken boys in terms of literacy scores. In literacy, girls in grades 6 (7) and higher had an average score that is nearly the same as the average for boys in grades 6 (7) and higher. The average unweighted score for in-school boys in numeracy was 58.3 percent, which is 6.7 percentage points above in-school, cohort girls (as compared with a gap of 8.4 percentage points at the baseline); the average score for boys in literacy was 63.1 percent, which is 3.6 percentage points above girls (as compared with a gap of 4.1 percentage points at the baseline). For reference, the midline sample of in-school boys and girls is composed of: n=163 boys; n=757 girls).<sup>71</sup> For numeracy, boys' average scores are higher than girls' scores by a statistically significant margin.<sup>72</sup> For literacy, boys' average scores are higher than girls', but the difference is not

<sup>71</sup> Note that the boys included in this analysis are those who have equivalent grade-levels to the girls to whom they are being compared. Thus, boys in the midline sample who reported being in secondary school grades are not included in this analysis.

<sup>72</sup> In a regression of numeracy score predicted by gender (and controlling for clustering at the school level), p=0.002.

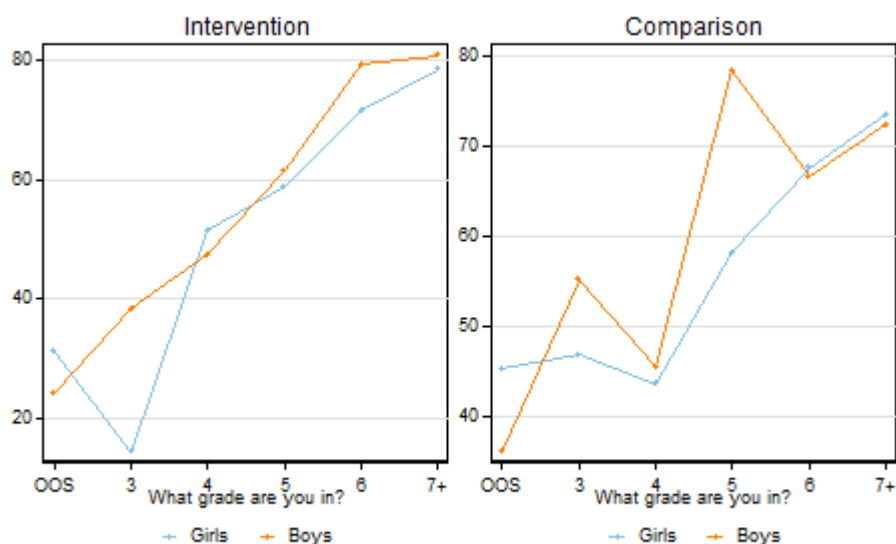


statistically significant, which is reflective of the fact that grade 3 (4) and grade 6+ (7+) girls have closed the gap with boys in literacy (see graphs below for a visualization of this finding).<sup>73</sup>

The panel of graphs below presents boys' literacy assessment scores (as percentages) alongside girls' assessment scores, by grade, including out-of-school boys and girls, and separated by intervention versus comparison. Boys' and girls' learning trajectories are no longer parallel (as they were at the baseline), which is predominantly a result of the fact that girls' learning at grade 3 (4) in the intervention group has now surpassed boys' learning at the same grade level. On the other hand, boys in the comparison group still outperform comparison girls consistently, albeit with a strong inflection point in boys' literacy from grade 3 (4) to grade 4 (5), which appears to be a result of uncharacteristically high performance of the subsample grade 4 (5) boys in the comparison group.

Using these graphs to reflect on the unexpected findings related to grade 3 (3) girls' performance, it is clear that grade 3 (3) girls in the intervention group under-performed relative to boys in the same grades and attending the same schools. On its own, this finding would suggest that the underperformance of grade 3 (3) girls may not be explained by school-level factors, or at least can only be attributed to school-level factors that affect boys differently from girls. On the other hand, the unexpectedly high performance by grade 3 (3) girls in the comparison group is mirrored by comparison boys in grade 3 (3).

*Figure 7: Comparison of boys' and girls' literacy scores, by grade and intervention status*



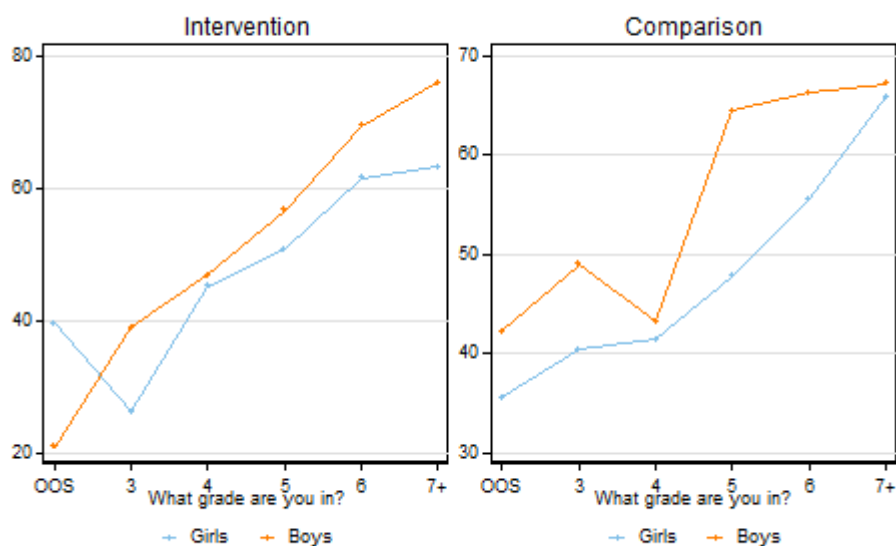
The panel of graphs below presents boys' numeracy assessment scores (as percentages) alongside girls' assessment scores, by grade, including out-of-school boys and girls, and separated by intervention versus comparison. All of the trends observed with regard to literacy scores above are also reflected in numeracy scores below, albeit to less dramatic extremes. On average, grade 3 (4) girls in the intervention

<sup>73</sup> In a regression of literacy score predicted by gender (and controlling for clustering at the school level),  $p=0.231$ .

group are performing at nearly the same level as boys in the intervention group. The gap between girls and boys in grade 3 (4) is somewhat wider for the comparison group.

Reflecting on the question of grade 3 (3) performance, the findings on numeracy are similar to those for literacy and similarly suggest that grade 3 (3) girls in the intervention group underperformed as compared with their peers in other grades as well as when compared with boys in the same grade. At the same time, comparison boys and girls in grade 3 (3) performed higher than what might be expected when compared with their peers in other grades.

Figure 8: Comparison of boys' and girls' numeracy scores, by grade and intervention status



The tables below summarise literacy and numeracy learning for boys by grade-level and by intervention versus comparison groups in analogous fashion to the summary presented for girls' learning above. Learning is positively correlated with grade-level, and increases monotonically by grade for the intervention group. In the comparison group, grade 3 (3) learners and grade 4 (5) learners are performing far above the average that we might expect based on the performance of their peers at the same grade-level in the intervention group. Similarly, grade 3 (3) and grade 4 (5) learners in the comparison group are performing above the averages for their peers in adjacent higher grades in the comparison group – i.e. comparison boys in grade 3 (3) are, on average, attaining higher literacy and numeracy scores than boys in grade 3 (4); and comparison boys in grade 4 (5) are, on average, attaining higher literacy scores than boys in grade 5 (6).

Table 27: Literacy (EGRA/SeGRA)

Grade	Intervention Group Mean	Comparison Group Mean	Standard Deviation in the intervention group

<b>OOS</b>	25.7	36.0	22.9
<b>Grade 3 (3)</b>	38.3	55.1	28.7
<b>Grade 3 (4)</b>	47.3	45.5	27.8
<b>Grade 4 (5)</b>	61.4	78.3	31.7
<b>Grade 5 (6)</b>	79.2	66.5	16.2
<b>Grade 6+ (7+)</b>	80.7	72.3	18.2
<b>Overall</b>	56.3	62.5	31.7

*Table 28: Numeracy (EGMA/SeGMA)*

<b>Grade</b>	<b>Intervention Group Mean</b>	<b>Comparison Group Mean</b>	<b>Standard Deviation in the intervention group</b>
<b>OOS</b>	22.3	42.2	19.8
<b>Grade 3 (3)</b>	38.9	49.0	26.3
<b>Grade 3 (4)</b>	46.8	43.2	22.1
<b>Grade 4 (5)</b>	56.7	64.4	20.9
<b>Grade 5 (6)</b>	69.4	66.2	18.7
<b>Grade 6+ (7+)</b>	75.9	67.2	23.6
<b>Overall</b>	52.3	57.8	27.9

Using boys' performance to reflect on cohort-girls' performance above, the unexpectedly high average scores for comparison-group boys in grade 3 (3) are results that parallel findings for comparison-group girls in grade 3 (3) above. The fact that both boys' and girls' scores exhibit this same trend suggests that the explanation for this finding may lie at the school or regional level, rather than at the individual level. This hypothesis will be explored further in the sections on intermediate outcomes (such as attendance and teaching quality) that tend to vary at the school-level and that may present an explanation for why grade 3 (3) learners in the comparison group have tended to perform above their expected level.

To further explore trends in boys' learning, the tables below present boys' literacy and numeracy scores by grade level, intervention versus comparison, and by baseline versus midline. These comparisons are structured in the same manner as the analogous tables for girls above, with the caveat that the sample of boys (from baseline to midline) is not a panel, but rather is a repeated cross-sectional sample.<sup>74</sup> Given the non-panel nature of the boys' sample, the table has been constructed such that the grade level shown in the **Cohort** column represents the grade of midline boys, while the baseline boys who form the basis for comparison are those who were one grade-level lower than the midline boys with whom they are being

<sup>74</sup> The non-panel nature of boys' data is important to note because difference-in-differences analysis relies on the parallel-paths assumption which does not hold with a repeated cross-sectional sample. Thus, in the above analysis of changes over time in boys' learning scores, it will be impossible to know whether observed changes over time are a result of project interventions (or other substantive causes that have altered learning trajectories over time), or are merely result of unobserved heterogeneity in the baseline versus midline samples.

compared. For example, boys in grade 3 (4) at midline are being compared with boys who were in grade 3 (3) at baseline, just as girls who were in grade 3 (4) at midline were compared to themselves at the baseline (with their baseline grade, in the majority of cases, being one grade lower than their midline grade).<sup>75</sup>

Boys' literacy scores exhibit similar trends to those of girls', with boys at each grade level showing improvement since baseline, but with the net amount of improvement since baseline being greater for comparison boys than for intervention boys.<sup>76</sup> In particular, grade 4 (5) boys in the comparison group at the midline have performed far above what would be expected given the performance of comparison boys at other grade-levels, as well as comparing with boys from the baseline sample. The result is that the difference from baseline to midline for comparison boys in grade 4 (5) is the largest observed (+34.7 percentage points).

Grade	Baseline Literacy Intervention	Midline Literacy Intervention	Difference Baseline to Midline (Intervention)	Baseline Literacy Comparison	Midline Literacy Comparison	Difference Baseline to Midline (Comparison)	Difference in Difference (Intervention Diff – Comparison Diff)
OOS	24.6	25.7	1.1	33.2	36.0	2.8	-1.7
Grade 3 (4)	37.6	47.3	9.8	32.2	45.5	13.3	-3.5
Grade 4 (5)	46.6	61.4	14.8	43.7	78.3	34.7	-19.9
Grade 5 (6)	58.4	79.2	20.8	53.9	66.5	12.7	8.1
Grade 6+ (7+)	64.3	80.7	16.4	58.8	72.3	13.5	2.9
Overall	40.9	56.3	15.4	42.7	62.5	19.8	-4.4

Turning to numeracy scores, boys at each grade level show improvement since the baseline, and the net improvement in numeracy is greater for comparison boys than for intervention boys.<sup>77</sup> As with literacy scores, grade-5 boys in the comparison group have the largest increase in numeracy score from baseline to midline.

Grade	Baseline Literacy Intervention	Midline Literacy Intervention	Difference Baseline to Midline (Intervention)	Baseline Literacy Comparison	Midline Literacy Comparison	Difference Baseline to Midline (Comparison)	Difference in Difference (Intervention Diff – Comparison Diff)
OOS	27.0	22.3	-4.7	32.5	42.2	9.7	-14.3

<sup>75</sup> Note that grade 3 (3) boys at the midline are omitted from this analysis because there are no grade 2 boys from the midline whose scores can form the appropriate basis for comparison.

<sup>76</sup> In a cluster-robust regression excluding OOS boys, the difference-in-differences estimate for literacy has a negative coefficient and  $p = 0.363$ .

<sup>77</sup> In a cluster-robust regression excluding OOS boys, the difference-in-differences estimate for numeracy has a negative coefficient and  $p = 0.720$ .

<b>Grade 3 (4)</b>	33.8	46.8	13.0	36.5	43.2	6.7	6.3
<b>Grade 4 (5)</b>	52.1	56.7	4.7	39.8	64.4	24.7	-20.0
<b>Grade 5 (6)</b>	53.0	69.4	16.4	52.6	66.2	13.6	2.8
<b>Grade 6+ (7+)</b>	58.4	75.9	17.5	58.3	67.2	8.9	8.6
<b>Overall</b>	40.0	52.3	12.3	42.2	57.8	15.6	-3.2

Across boys' literacy and numeracy, grade 4 (5) is the primary source of negative difference-in-differences, with comparison boys in grade 4 (5) having experienced a far greater rate of improvement since baseline than intervention boys in grade 4 (5). The extremely high performance of comparison boys in grade 4 (5) is not paralleled in girls' scores, and appears to be an anomaly that may simply be a result of the very small subsample size of grade 4 (5) boys at the midline (N = 34), as well as the fact that the sample is a repeated cross-section (not a panel).

The tables on the following page present foundational skills gaps analysis for boys. The primary finding is that skill gaps for boys are essentially the same as those for girls, and have remained parallel to girls' skill gaps since the baseline. More specifically, the performance gap between boys' and girls' scores was substantial at the baseline, but their primary skill gaps were in the same places. At midline, the overall performance gap has closed somewhat, and skill gaps remain parallel, with the proportion of proficient learners in a given category for boys being roughly equivalent in most subtasks to the proportion of proficient learners in a given category for girls.

As with girls, the primary skill gap for boys in numeracy is between basic subtraction and more advanced addition (subtask 3 to subtask 4), and the primary skill gap for boys in literacy is between reading words and basic reading comprehension (subtask 1 to subtask 2). The same caveat applies to analysis of reading comprehension subtask 2 here, given the fact that this subtask was significantly more difficult at midline than at baseline. Nonetheless, it is clear for boys (as for girls) that there have been major improvements in the ability to read words as well as reading fluency, but there have not been corresponding improvements in reading comprehension. A more detailed analysis of girls' skill gaps has been presented above and because boys' skill gaps are nearly identical, it is not necessary to replicate this analysis for boys.

In general, skill development between girls and boys is currently parallel, and girls appear to be on a trajectory to close the overall performance gap over time. At this level, there is preliminary evidence that the project is helping girls to close the gender-based performance gap that existed at baseline, and evidence on skill development suggests that this trend will continue. However, these findings should also be interpreted in light of the fact that the duration of the intervention has been less than one year.

Although the qualitative data largely supports the quantitative finding that there is still progress to be made in girls' learning, there is also evidence in the qualitative data to suggest that girls have actually surpassed boys in some schools. Girls were still commonly described as shy, but some respondents highlighted that girls are now outperforming boys: "Yes, there are changes in the situation over the past year. There is more awareness and support for girls' education. You'll notice that girls perform better in school than boys."<sup>78</sup> Another respondent attributed this change to the project, suggesting it has been successful in improving girls' education and performance at school: "I believe the improvement of the quality of the girls doubled compared to the boys, as I have told you, and this was a result of the hard work from CARE International, teachers, school administration and the parents of the students."<sup>79</sup> Although this finding that girls are performing better in school is positive, it also suggests that boys' inclusion should be an important consideration for the project moving forward, a topic which will be discussed further in later sections of the report.

Additionally, boys appear in some cases to be less motivated than girls, and are often characterized as being more concerned with football or chewing khat than applying themselves in school. When asked about the differences between boys and girls in school, one teacher explained, "Honestly, even though last year I was not here, if you track it, last year's girls were doing well and boys are not well motivated. They hesitate to make efforts in education."<sup>80</sup> A mother provided further evidence of this, explaining that her daughters are more serious about their studies than her sons: "Girls and boys are not same. I have four girls, all of them they are studying at the university and I have two boys who dropout from the primary school and the secondary school. So I believe that girls are better when it comes the education."<sup>81</sup>

Misinterpretation of messaging around the importance of girls' education may be partially responsible for this lack of motivation observed among boys. For example, it appears that girls are at times encouraged at the expense of boys. One teacher explains the following method which is used to motivate girls in the community: "Every morning when the students are in the line, we encourage the girls and we tell them to encourage other girls to come to the school. We tell them that boys are busy on chewing khat and smoking cigarettes, so the girls are required to take that role and help their families. Now, we are planning to teach the girls who work in the tea shops."<sup>82</sup>

These qualitative findings do not appear to be representative of the norm, as they were not mentioned with enough frequency to suggest that boys are being systematically excluded or disadvantaged. However, the findings are worth noting to ensure that the project remains sensitive to and continues to monitor the effects of girls' education programming on boys.

---

<sup>78</sup> FGD – Teachers

<sup>79</sup> KII – MoE

<sup>80</sup> FGD – Teachers

<sup>81</sup> FGD – CEC

<sup>82</sup> FGD – Teachers

Table 29: Boys' foundational numeracy skills gaps

Categories	Subtask 1	Subtask 2	Subtask 3	Subtask 4	Subtask 5	Subtask 6	Subtask 7	Subtask 8	Subtask 9	Subtask 10	Subtask 11
	Missing Number	Addition (Level 1)	Subtraction (Level 1)	Addition (Level 2)	Subtraction (Level 2)	Word problems (add/subtract)	Multiplication (Level 1)	Multiplication (Level 2)	Division (Level 1)	Division (Level 2)	Word problems (mult/div)
<b>Non-learner 0%</b>	6.6 (-1.8)	7.5 (-3.2)	10.4 (-22.3)	21.7 (-17.1)	26.4 (-23.1)	17.9 (-18.5)	32.1 (-20.7)	68.9 (-6.4)	53.8 (-16.3)	75.5 (-7.2)	58.5 (-9.7)
<b>Emergent learner 1%-40%</b>	29.2 (1.7)	2.8 (-1.4)	3.8 (-1.4)	17 (0.2)	9.4 (-5.5)	3.8 (-3.7)	5.7 (-5.1)	11.3 (1.5)	11.3 (1.5)	8.5 (1)	0 (0)
<b>Established learner 41%-80%</b>	34.9 (8.7)	19.8 (2.1)	11.3 (0.6)	23.6 (7.2)	26.4 (8.7)	30.2 (6.4)	33 (18.5)	7.5 (3.3)	13.2 (2.9)	7.5 (5.2)	14.2 (2)
<b>Proficient learner 81%-100%</b>	29.2 (-8.6)	69.8 (2.5)	74.5 (23.1)	37.7 (9.7)	37.7 (20)	48.1 (15.9)	29.2 (7.3)	12.3 (1.5)	21.7 (11.9)	8.5 (1)	27.4 (7.7)
	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0



Table 30: Aggregate DID analysis of boys' mean numeracy scores by subtask

Categories	Subtask 1 Missing Number	Subtask 2 Addition (Level 1)	Subtask 3 Subtraction (Level 1)	Subtask 4 Addition (Level 2)	Subtask 5 Subtraction (Level 2)	Subtask 6 Word problems (add/subtract)	Subtask 7 Multiplication (Level 1)	Subtask 8 Multiplication (Level 2)	Subtask 9 Division (Level 1)	Subtask 10 Division (Level 2)	Subtask 11 Word problems (mult/div)
<b>Intervention Baseline</b>	62.2	79.3	59.1	45.3	35.3	50.2	36.0	16.3	19.5	10.8	25.7
<b>Intervention Midline 1</b>	59.5	83.8	82.9	60.6	60.2	69.6	54.2	20.6	33.6	16.0	34.4
<b>Comparison Baseline</b>	64.1	79.5	70.1	53.1	40.8	53.5	38.6	14.8	21.4	8.8	19.8
<b>Comparison Midline 1</b>	61.8	88.0	90.9	71.4	62.0	75.3	61.8	19.7	43.0	18.0	43.7
<b>DiD</b>	-0.4	-4.0	3.1	-3.0	3.6	-2.4	-5.0	-0.6	-7.6	-4.0	-15.1

Table 31: Boys' foundational literacy skills gaps

Categories	Somali ST1 Reading Words	Somali ST2 Reading Comp (easy)	Somali ST3 Reading Comp (medium)	Somali ST4 Reading Fluency	Somali ST5 Reading Comp (difficult)	Somali ST6 Writing (fill blank)	Somali ST7 Writing (negative form)	Somali ST8 Writing (future tense)
<b>Non-learner 0%</b>	3.8 (-12.2)	7.6 (-12.2)	13.9 (-14.8)	10.1 (-18.6)	22.8 (-24.4)	26.6 (-18.2)	35.4 (-19.7)	46.8 (-14)
<b>Emergent learner 1%-40%</b>	2.5 (-41.8)	7.6 (3.8)	8.9 (-2)	11.4 (-11.2)	15.2 (8.6)	13.9 (0.7)	2.5 (-4.5)	3.8 (-4.7)
<b>Established learner 41%-80%</b>	15.2 (-24)	29.1 (3.2)	45.6 (21)	46.8 (12.9)	41.8 (12.1)	24.1 (12.3)	15.2 (6.7)	12.7 (4.2)
<b>Proficient learner 81%-100%</b>	78.5 (78)	55.7 (5.2)	31.6 (-4.2)	31.6 (17)	20.3 (3.7)	35.4 (5.3)	46.8 (17.6)	36.7 (14.5)
	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Table 32: Aggregate DID analysis of boys' mean numeracy scores by subtask

Categories	Somali ST1 Reading Words	Somali ST2 Reading Comp (easy)	Somali ST3 Reading Comp (medium)	Somali ST4 Reading Fluency	Somali ST5 Reading Comp (difficult)	Somali ST6 Writing (fill blank)	Somali ST7 Writing (negative form)	Somali ST8 Writing (future tense)
<b>Intervention Baseline</b>	32.0	65.3	54.7	40.2	31.4	41.0	33.3	29.4
<b>Intervention Midline 1</b>	80.1	69.8	60.8	59.2	35.8	49.4	52.5	42.8
<b>Comparison Baseline</b>	32.1	68.9	55.2	37.8	37.3	42.5	37.3	30.7
<b>Comparison Midline 1</b>	85.3	75.9	63.6	64.4	50.3	56.1	57.8	46.4
<b>DiD</b>	-5.1	-2.6	-2.2	-7.6	-8.6	-5.3	-1.4	-2.4

The table below presents boys' learning outcomes for in-school boys disaggregated by relevant subgroups (as well as the score for out of school boys, presented separately below for comparison). The results of this subgroup analysis parallel baseline findings (for boys) and also parallel the subgroup analysis findings that will be presented below for girls. Out-of-school status remains the strongest predictor of lower than average learning outcomes, and is associated with a net decrease in literacy and numeracy scores since the baseline.

Table 33: Learning scores of boys' subgroups

	Average literacy score (aggregate)	Change in average literacy score since baseline	Average numeracy score (aggregate)	Change in average numeracy score since baseline	Number of observations for subgroup
<b>Characteristics:</b>					
All in-school boys	58.3	11.2	63.1	14.5	163
Out of school boys	27.7*	-1.8	28.5*	-0.1	22

\*Note, an asterisk indicates results that are statistically significant at the 95% confidence level (or higher) in a regression with cluster-robust standard errors.

Taken as a whole, the findings for boys' learning outcomes suggests that boys and girls are learning at increasingly similar rates and facing similar challenges in terms of the acquisition of new skills. Regional trends are similar between boys and girls, suggesting that the macro-level factors (such as school curriculum and funding) that potentially affect learning outcomes are affecting boys and girls in similar ways in terms of their learning.

### Grade Levels Achieved

This section shifts back to an analysis of girls' learning outcomes and reports the share of girls achieving each grade level of numeracy, based on an analysis of the school curricula or syllabi. Where skills were not specifically listed in either curriculum, the evaluation team chose the best subjective match possible.<sup>83</sup> Finally, the mapping of skills to grade-levels at baseline focused on English literacy and mathematics only, due to the lack of information on expected learning outcomes for Somali literacy.<sup>84</sup> Because English literacy was not evaluated at midline, the analysis below focuses exclusively on numeracy.

Table 34: Grade Level Standards for Numeracy

Grade Level Achieved	Numeracy Skills
1	<ul style="list-style-type: none"> <li>Number identification up to 99 (portion of subtask 1)</li> <li>Addition without carrying numbers (portion of subtask 2)</li> <li>Subtraction without borrowing (subtask 3)</li> </ul>
2	<ul style="list-style-type: none"> <li>Number identification up to 999 (portion of subtask 1)</li> <li>Addition carrying one number (portion of subtask 2)</li> <li>Addition with 3 digits, carrying up to 1 number (subtask 4)</li> <li>Subtraction carrying one number (portion of subtask 5)</li> <li>Addition and subtraction word problems with simple underlying arithmetic (subtask 6)</li> <li>Multiplication of 1-digit numbers (subtask 7)</li> </ul>

<sup>83</sup> As an example, neither curriculum available specifies when a child should learn to construct the negative form of a sentence. We consider this skill on par with the difficulty of constructing future tense sentences, which is a skill expected to be developed in Grade 6.

<sup>84</sup> It is also important to note that the available curricula are focused on primary-level outcomes. Given the evaluation's focus on students in grades 3-5 (3-6), this does not pose a problem for the baseline evaluation. However, a fuller understanding of grade level achievement at the endline may necessitate a deeper review of curricular materials – where available – for secondary schools.

	<ul style="list-style-type: none"> <li>• Division of 2-digit number by 1-digit number (subtask 9)</li> </ul>
3	<ul style="list-style-type: none"> <li>• Subtraction carrying two numbers (portion of subtask 5)</li> <li>• Multiplication of 2-digit numbers (subtask 8)</li> <li>• Word problems with simple multiplication and division (subtask 11)</li> </ul>
4	<ul style="list-style-type: none"> <li>• Division of 3-digit number by 2-digit number (subtask 10)</li> </ul>
5	N/A
6	N/A

The table above describes the standards developed by the evaluation team for grade level achievement in numeracy. In cases where no standard is described, the learning assessments utilised in the evaluation did not include a skill specific to that grade level. In cases in which a subtask is distributed across two grade levels, the table and coding of level achieved distinguishes between less and more difficult portions of the subtask, on an item-by-item basis. In order to achieve a given grade level, a student must achieve a score of approximately 80 percent or higher on subtasks (or relevant, grade-specific portions of a subtask) for that grade, and those for the preceding grades.<sup>85</sup>

In the table below, the achievement levels for cohort girls in the intervention group are provided alongside comparison cohort girls, whose levels are reported in parentheses to facilitate side-by-side comparison. Each cell presents the percentage of girls at a given midline grade-level (columns) who have achieved a given grade-level in terms of their performance on relevant subtasks (rows).

*Table 35: Numeracy grade level achieved by intervention (and comparison girls in parentheses), by Grade*

	Out-of-school	Grade 3 (3)	Grade 3 (4)	Grade 4 (5)	Grade 5 (6)	Grade 6+ (7+)
Grade 1 achieved	20.7 (32.4)	24.1 (36.8)	36.9 (27.2)	23.7 (42.2)	31.3 (27.4)	34.8 (30.3)
Grade 2 achieved	0 (0)	0 (0)	1.2 (2.4)	5.2 (0)	7.3 (8.3)	5.9 (11.3)
Grade 3 achieved	0 (0)	0 (0)	0 (1.5)	0 (0)	2.2 (2.2)	3.4 (0)
Grade 4 achieved	0 (4.1)	0 (0)	4.3 (0.8)	7 (1.9)	6.8 (5.5)	12.6 (7.7)

The primary finding at the baseline was that a very small proportion of girls had achieved levels of performance that would be expected given their grade-levels. At midline, there has been little change, albeit with a slight increase in the proportion of intervention girls who are in grade 3 (4) and are also performing at a grade-4 level (an increase from 2.5 percent at baseline to 4.3 percent at midline). The only other trend that has emerged since the baseline is a moderate level of bimodality in achievement levels, where girls who are progressing well in their studies have attained a grade-4 achievement level by

<sup>85</sup> Where there are many items in a subtask, we follow the 80 percent rule. Where there are four items, we allow one wrong answer (75 percent correct) without disqualifying a student from achievement of a grade level.

the time they are at or beyond grade 3 (4) in school, while their peers who are learning more slowly tend to be two or even three achievement levels lower (i.e. having a grade 1 or grade 2 achievement level, with a gap in grade 3 (3) in terms of achievement levels).

Using grade-level analysis as another means of reflecting on differential performance of grade 3 (3) girls in intervention versus comparison, it is clear that none of the grade 3 (3) girls in the sample are performing at their expected grade-level. For grade 3 (3) girls, the differentiating factor is the proportion who have attained a grade-1 achievement level. In keeping with the findings presented above, the proportion of grade 3 (3) comparison girls who have at least reached a grade-1 achievement level is 12.7 percentage points higher than the proportion of grade 3 (3) intervention girls (24.1 percent in the intervention group and 36.8 percent in the comparison group).

Comparing girls' achievement levels with boys' achievement levels, table below presents the numeracy achievement levels for boys in the intervention group alongside comparison boys, whose levels are reported in parentheses to facilitate side-by-side comparison.

*Table 36: Boys' numeracy grade level achieved by intervention (and comparison boys in parentheses), by Grade*

	Out-of-school	Grade 3 (3)	Grade 3 (4)	Grade 4 (5)	Grade 5 (6)	Grade 6+ (7+)
Grade 1 achieved	20 (16.7)	25 (50)	36.4 (41.7)	43.5 (50)	50 (42.9)	40.9 (36.8)
Grade 2 achieved	8 (0)	6.3 (7.1)	4.5 (0)	4.3 (7.1)	8.3 (14.3)	9.1 (10.5)
Grade 3 achieved	4 (0)	0 (0)	0 (0)	0 (7.1)	12.5 (0)	18.2 (5.3)
Grade 4 achieved	0 (0)	0 (0)	0 (8.3)	4.3 (0)	4.2 (7.1)	22.7 (10.5)

The main finding for boys (as with girls) is that nearly all learners have achievement levels that are well below their grade level. Also paralleling the findings for girls, a higher proportion of grade 3 (3) boys in the comparison have reached grade-1 and grade-2 achievement levels than grade 3 (3) boys in the intervention group.

The primary point of divergence between boys' and girls' achievement levels is at grade levels 5 to 6+ (6 to 7+), where the proportion of boys who have achieved grade 3 and grade 4 achievement levels far outstrips the proportion of girls whose scores place them at the same achievement level. This finding is in keeping with the comparative analysis of boys' versus girls' learning scores presented above: girls have closed the gap with boys at the level of grade 3 (4), but the performance gap between boys and girls widens at higher grade-levels.

## 5.2. Subgroup analysis of Learning Outcomes

This section presents an analysis of learning outcomes by key subgroups of the population of cohort girls, as well as an analysis of potential barriers to learning. The section on boys' learning outcomes has already addressed the issue of contrasts between girls and boys, so this section focuses on key differences within the sample of girls. As at the baseline, there were very few critical subgroupings that helped to identify girls who are likely to score significantly lower than their peers. As at the baseline, OOS status, and several proxies of pastoralism are significant predictors of lower than average learning

outcomes. The analysis of barriers to learning allows for the confirmation of several critical barriers that are consistent predictors of lower learning outcomes and that have persisted since the baseline. In particular, measures of school and teaching quality (as reported by the caretaker) are the strongest predictors of learning outcomes.

The table below summarises midline learning scores by literacy and numeracy (and changes in average scores from baseline to midline) for each of the major subgroups within the sample of 766 in-school cohort girls. The subsample of in-school girls is used here in order to focus on factors that potentially explain differences in learning outcomes among girls who are attending school and to maintain comparability with the presentation of these results at baseline. Out-of-school girls are also presented as a separate subgroup for analysis in the final row of the table below.

	Average literacy score (aggregate)	Change in average literacy score since baseline	Average numeracy score (aggregate)	Change in average numeracy score since baseline	Number of observations for subgroup
<b>Characteristics:</b>					
All in-school girls	51.1	11.8	58.8	16.8	766
Living without both parents	50.8	12.0	60.5	18.0	72
Mother tongue different to LOI	58.5	0.6	65.5	19.1	44
<b>Disability</b>					
Vision impairment	29.8	-5.0	52.0*	7.5	3
Hearing impairment	0.0	0.0	0.0	0.0	0
Mobility impairment	50.9*	-13.2	67.9	58.5	1
Cognitive impairment	0.0	0.0	0.0	0.0	0
Self-care impairment	0.0	0.0	0.0	0.0	0
Communication impairment	52.3*	30.9	79.2	26.0	1
Mental health impairment	56.1	12.1	62.9	13.8	94
Anxious	56.0	12.7	63.0	13.3	84
Depressed	55.3	6.3	60.4	10.1	60
Any disability	55.3	11.6	62.6	13.6	97
<b>HOH and Carer Characteristics</b>					
HOH no wage-earning occupation	52.0	11.6	60.3	19.0	414
HOH no education	50.5	12.7	57.6	18.7	385
HOH female	55.8*	14.7	61.8*	19.6	346
HOH Pastoralist	47.1*	13.3	49.6	9.1	54
Carer no education	50.3	12.2	58.7	20.2	427
<b>Household Assets</b>					
Owns camels	52.8	15.8	54.8	14.6	81

Owns medium-sized livestock	50.3*	12.2	57.4	16.3	523
Owns small livestock	47.7	11.9	50.5	10.1	40
Owns mobile phone	50.4	10.5	58.2	16.2	668
Access to water reservoir/storage	52.7	9.5	60.5	15.8	463
Owns land	50.7	12.8	58.7	17.0	531
<b>Poverty</b>					
House is informal/temporary structure	48.7	10.9	53.8	17.3	50
Gone to sleep hungry most days	49.6	19.1	50.3	15.7	28
Gone without enough clean water most days	49.6	13.9	54.2	16.0	78
Gone without medicines or medical intervention most days	54.1	15.0	58.5	17.9	147
Gone without cash income most days	53.1	12.6	59.0	15.8	153
<b>Migration</b>					
Displaced or moved in past 12 months	48.5	10.2	65.1	-8.0	4
Household migrates seasonally	50.1	26.4	46.9	16.0	11
<b>Other</b>					
High chore burden (whole day spent on chores)	49.7	4.0	58.9	2.1	66
Married	51.4	7.2	48.5	3.1	6
Out of school girls	37.1*	3.8	40.1*	5.8	41

\*Note, an asterisk indicates results that are statistically significant at the 95% confidence level (or higher) in a regression with cluster-robust standard errors.

As at the baseline, learning outcomes were not significantly correlated with most subgroup types. The most notable shift from the baseline in terms of subgroups that are strong predictors of learning outcomes is that girls with a female head of household now tend to perform significantly above average in both literacy and numeracy. Among girls who are conventionally considered at very high risk of dropping out of school, being married, having a child, as well as living with her husband or parents-in-law are also not significant predictors of learning outcomes. It is also worth noting that proxies for poverty and migration status are also not significant predictors of lower learning outcomes at the midline, despite the fact that these were significant predictors of lower learning outcomes at the baseline.

A few of the subgroups above have statistically significant correlations with learning outcomes that have been consistent since the baseline.

The qualitative data suggest that conflict during the past year may help to explain some variation in learning outcomes. Focus group respondents from some locations reported local bouts of violent conflict in the recent past. In some cases, the effects were severe – students in some communities lost almost an entire academic year – and although the situation has improved in certain areas, other areas are still recovering. CEC members at a conflict-affected school explained that clan conflict last year forced community members to move out of the area. In the past, soldiers used the school as a military base, and a number of classrooms that were built by CARE are reportedly still being occupied by soldiers: “Soldiers



are still occupying two rooms built by CARE for the school and the Ministry has not removed them yet.”<sup>86</sup> Understanding that learners in conflict-affected parts may have had attenuated learning as a result of conflict-related school closures and reduced learning time, it is worth considering whether this particular regional difference may help to explain the differential grade 3 (3) learning outcomes noted earlier. The table below presents the composition of the sample of grade 3 (3) learners by intervention versus comparison and by region.

Due to uneven geographic distribution, a much larger proportion of the girls in the intervention grade 3 may have been affected by conflict and resulting school-closures. This finding does not help to explain why girls in the comparison group have performed so much better than average, but it does help to explain the extreme under-performance of grade 3 (3) girls in the intervention group.

General disability status is not a strong predictor of lower learning outcomes, with the exception that girls who have vision impairments have lower than average scores in both literacy and numeracy (with the correlation being statistically significant in the case of numeracy). Due to the change in how disabilities were measured and coded at baseline versus at midline, longitudinal comparisons are not valid, since the observed changes over time are most likely driven by changes in coding (hence changes in sub-sample membership). All results presented here are based on the full midline sample with the new disabilities coding parameters, in order to create the most straightforward and intuitive setup for longitudinal comparison moving forward. At the baseline, disability status, especially indicators of mental health impairments and trauma were strong predictors of lower than average learning outcomes in both literacy and numeracy (with the correlation between mental health disability and lower literacy being statistically significant at baseline). In contrast, at the midline girls with mental health impairments as well as girls with signs of anxiety or depression are all performing at or above the sample average. Due to the comparability problems noted above, we cannot conclude that the situation of girls with mental health impairments has in fact changed and improved. Rather, we can observe that, based on midline learning scores and midline coding criterion, girls with mental health and trauma-related impairments do not appear to be marginal in the ways that were emphasized as part of the baseline report findings.

Turning to physical disabilities, the finding that disability status is no longer predictive of lower learning outcomes is at least partly a result of reduction in the number of girls in the sample who have physical disabilities due to the non-tracking of OOSG in this evaluation round. Although overall disability status does not predict learning outcomes that are significantly different from average, vision impairments remain (since the baseline) a strong predictor of lower than average learning outcomes, with numeracy outcomes being lower to a statistically significant degree for girls with vision impairments.

At the baseline, a pastoralist lifestyle was identified as a potential predictor of lower learning outcomes, and the evidence at the midline provides further support for this hypothesis. The effects of pastoralism on learning can be assessed through both the stated profession of the head of household, as well as the status of a household as owning small or medium-sized livestock. The most direct measure of pastoralism is if the head of household's occupation is reported to be pastoralism or animal herding, and at the midline belonging to this subgroup is a predictor of lower learning outcomes in both literacy and numeracy (with literacy scores for girls belonging to pastoralist households being lower than average to a statistically significant degree). At baseline, ownership of medium-sized livestock (taken as another proxy for pastoralism) was also a predictor of lower than average learning outcomes, and this finding persists at the midline as well, with literacy scores being lower to a statistically significant degree among girls

---

<sup>86</sup> FGD – CEC members

belonging to households that own medium-sized livestock. As at the baseline, seasonal household migration is not significantly predictive of lower learning outcomes, and nor are other potentially related variables such as higher chore-burden. At baseline, living in an informal or temporary structure was a significant predictor of lower learning outcomes, and at the midline the scores of girls living under such conditions are also lower than average, albeit not to a statistically significant degree. While the evidence is somewhat mixed, most indicators point to the conclusion that pastoralism is still a strong predictor of lower than average learning outcomes at the midline.

The qualitative data on the effects of drought on pastoralist households helps to establish why, despite project efforts to address the needs of learners in pastoralist families, girls belonging to pastoralist households continue to perform below average at the midline. While drought in Somalia has been widespread and has affected all households to some extent, pastoralist families have experienced the effects of drought most acutely because many have lost significant portions of their livestock (hence livelihoods) and been forced in some cases to move into new areas in search of assistance. As one community member explains, “All of the community members have struggled from bad droughts. This community is 100% pastoralist, and the droughts killed off all of our livestock, leaving us with nothing.”<sup>87</sup> The resulting internal displacement has led to an increase in enrolment in certain areas, particularly cities, and a decrease in enrolment in other predominantly rural areas. Overall, migrants and residents alike have become more vulnerable as a result of the drought -- either directly through loss of livestock or indirectly as communities have struggled to absorb and support new arrivals.

The qualitative findings presented above suggest that financial hardship is the primary mechanism through which drought affects the education of girls in pastoralist households. From a short-term economic standpoint, families who have lost their main source of livelihood because of drought may be unable to afford school fees or lose a potential income-earner for the family by sending their children to school. When one mother was asked whether she will be able to continue sending her girls to school, she explained that most families “won’t be able to afford it as they have lost their livestock.”<sup>88</sup> Clearly, economic hardship resulting from the drought made it difficult for some families to afford keeping their children in school consistently (when can help to explain dropouts as well as reduced attendance), and economic hardship also can provide incentives to keep girls out of school so that they can help support the household economy.<sup>89</sup>

As an important nuance to the discussion above, drought-induced economic hardship appears to have had the greatest effect on girls’ learning among pastoralist families, but the broader effects of poverty appear to have been reduced since the time of the baseline. The mere fact that a household shows evidence of experiencing economic hardship (as measured through poverty proxies) is not, in and of itself, a strong predictor of lower learning outcomes at the midline, even though it was a fairly consistent and significant predictor of lower learning outcomes at the baseline. Qualitative evidence suggests that this shift in the salience of general poverty as a predictor of learning outcomes may be a result of the fact that the cost of schooling, which was cited as a significant barrier to initial and continued enrolment at the baseline, is in some instances no longer as prohibitive of a factor for vulnerable families at the midline.

When discussing the financial aspects of keeping girls in school, many midline respondents specifically mentioned the positive impact CARE International’s intervention has had on community mobilization around girls’ education in this sense. In particular, CECs in many areas have focused on fundraising and

---

<sup>87</sup> FGD – CEC members

<sup>88</sup> FGD – Mothers

<sup>89</sup> FGD – Mothers

raising awareness among community members, and teaching quality has improved as CARE International has provided training to teachers and support aimed at building the capacity of local education sector stakeholders. A CEC member explains the change seen within the community: “The number of students is increasing compared to last year. The awareness contributes to students being able to learn well. The quality of girls’ education is increased by the support they receive, for example the free education offered by CARE International. [...] The students are receiving encouragement, and the girls who have dropped out are coming back to school and that is because of the awareness.”<sup>90</sup>

Finally, out-of-school status remains a strong predictor of lower learning outcomes (since the baseline). Out-of-school girls have literacy and numeracy scores that are lower than the averages for in-school girls to a statistically significant degree. At the midline, all OOS girls in the sample dropped out during the previous year, and thus the extremely low learning outcomes among this subgroup are a direct illustration of the important correlation (likely a result of reciprocal causation) between transition and learning outcomes.

The table below presents potential barriers to learning, with average learning scores for sets of learners who reported having encountered one of the barriers below (as well as changes in average scores from baseline to midline for girls who reported experiencing a given barrier). These different barriers are relatively common, as can be seen from the fairly large subgroup sizes reported below. Many of the same barriers that were significant predictors of lower learning outcomes at the baselines are also predictors of lower outcomes at the midline as well.

*Table 9: Learning scores of key barriers*

	Average literacy score (aggregate)	Change in average literacy score since baseline	Average numeracy score (aggregate)	Change in average numeracy score since baseline	Number of observations for subgroup
<b>Barriers:</b>					
All in-school girls	51.1	11.8	58.8	16.8	766
<b>School Infrastructure</b>					
Difficult to move around school	60.4	20.5	62.9*	22.0	154
Doesn't use drinking water facilities	56.4	11.9	61.8	17.8	127
Doesn't use toilet at school	56.2	14.0	61.3*	15.5	185
Doesn't use areas where children play/socialise	51.1	11.2	57.1	17.9	327
<b>School Resources</b>					
No computers at school	50.6	11.7	58.6*	16.5	720
School does not have learning materials	51.6	11.9	59.6	16.4	186
Not enough seats for children at school	50.9	15.9	62.3	25.6	115
<b>Teaching Quality</b>					

<sup>90</sup> FGD – CEC members

Disagrees teachers make them feel welcome	49.0	18.2	54.6	14.1	34
Agrees that they are afraid of teacher	51.1	14.3	58.7	18.4	592
Agrees teachers treat boys and girls differently in the classroom	53.8	13.2	60.7*	18.7	332
Agrees teacher is often absent from class	52.2	15.3	56.6	10.8	166
Teacher punishes students who get things wrong	51.4	12.7	58.7	17.1	635
Teacher uses corporal punishment	49.7*	11.4	57.5	15.4	231
Carer says principal performance is poor	35.1*	-2.2	39.1*	-7.5	23
Carer says teaching at school is poor	37.0*	11.2	41.6*	14.4	18
<b>Gender Equity</b>					
Teacher targets questions by gender	45.8*	6.8	51.5	13.1	74
Teacher targets difficulty of questions by gender	46.7*	10.6	48.8	14.1	101
<b>Other Barriers</b>					
Agrees she has no choice in schooling decisions	51.5	12.8	59.3	17.3	681
Over 30-minute travel time to school	43.4*	11.2	44.5*	11.9	47
Feels unsafe on way to school	62.3	13.6	69.0*	18.0	23
Feels unsafe at school	47.5	14.5	51.8	9.7	13

\*Note, an asterisk indicates results that are statistically significant at the 95% confidence level (or higher) in a regression with cluster-robust standard errors.

Proceeding from the top, down, none of the indicators of school infrastructure had a consistent effect on girls' learning outcomes. This finding at the midline is consistent with the baseline findings related to barriers.

In keeping with baseline findings, students in better-resourced schools at the midline (i.e. those with computers) consistently outperformed those in schools without computers. The table above shows the obverse of this relationship, namely that attending a school with no computers is a consistently negative predictor of learning outcomes (vis-à-vis students who attend schools that do have computers). The result for numeracy outcomes is statistically significant at the midline, and the result for literacy outcomes is not. Note that this result is not about the effect of computers on learning in the classroom per se, but rather about computers serving as a proxy for how well-resourced a school is in general.

In contrast to baseline findings, having adequate learning materials is no longer a significant predictor of lower literacy outcomes.

Turning to indicators of teaching quality, poor principal performance and poor teaching (as reported by the caretaker) were both predictors of lower than average learning outcomes for literacy and numeracy at the baseline, and both of these barriers remain significant predictors of lower learning outcomes in both

literacy and numeracy at the midline. Also in keeping with baseline findings, use of corporal punishment is associated with lower than average learning outcomes and is a statistically significant predictor of literacy outcomes. The analysis of teaching quality as an intermediate outcome will explore these relationships in greater detail, but for now it will suffice to observe that there has been little improvement in teaching quality in the intervention group over and above improvements in the comparison group, which may help to explain why learning in the intervention group has also not improved significantly over and above the comparison group.

Qualitative data helps to underline the salience of teaching quality by underlining the disproportionate influence a single teacher can have (whether high quality or low quality) in schools with unfavourable teacher-student ratios. In one school, a girl explained during the qualitative interview that although there used to be three teachers in her school, there is now only one teacher because one became sick and another was kicked out of the school. A disabled girl from another school explained how she does not feel comfortable asking a certain teacher for help: “Yes, I ask the teachers for help, except for one of them. If you tell that teacher that you did not understand the lesson and if he could please repeat it, he says it is not any of his business.”<sup>91</sup>

In keeping with baseline findings (and with the findings above related to teacher quality in general), gender inequality in the classroom is also predictive of lower learning outcomes in both literacy and numeracy at the midline, and this relationship is statistically significant for literacy at the midline. Even though the gender gap in learning appears to be closing, it is important to bear in mind that problems with unequal treatment of boys and girls by teachers still exist in the classroom, and that these inequalities still influence learning in significant ways. In light of the fact that girls already struggle with shyness in the classroom and that teacher numbers are already so low in some schools, a school that has even one teacher who does a poor job of encouraging girls and addressing their needs could potentially be registering as having lower learning outcomes solely because girls are afraid to speak up when they do not understand lessons. A teacher explained girls’ shyness in the classroom in the following way: “When you have finished teaching them a lesson and you ask the students if they understood, the girls always say ‘yes,’ but they did not understand at all. In fact, they are too shy to say out loud that they did not understand. The boy will tell you that he did not understand the lesson straight away.”<sup>92</sup> This teacher’s testimony about girls’ shyness certainly suggests that teachers are aware of gendered differences in their classrooms, but it also seems to convey an attitude that problematises the girls’ behaviour, rather than conveying an understanding of how to better address girls’ needs in a society that tends to discourage girls from openly speaking their minds and asserting their needs.

Finally, a long travel time to school (over 30 minutes) is predictive of significantly lower learning outcomes in both literacy and numeracy. This barrier is fairly exceptional (with only about 6 percent of girls reporting having such a long trip), but it is nonetheless a significant barrier that has persisted since the baseline.

#### *Understanding the Performance Gap in Grade 3 (3) Intervention versus Comparison Girls*

Before concluding this analysis of learning outcomes for cohort girls, it is worth returning to the question of why grade 3 (3) intervention girls scored much lower than might be expected while grade 3 (3) comparison girls scored so much better than expected. The subgroups and barriers analysis above suggests some important household and school-level factors that can explain uneven learning outcomes, but these factors do not vary significantly between grade 3 (3) girls in the intervention versus the comparison group, and thus these factors cannot explain the divergent outcomes observed. The primary

---

<sup>91</sup> KII – Girls with disability

<sup>92</sup> FGD – Teachers

explanation that arises from the quantitative analysis and that is also corroborated by the qualitative analysis is the regional differences noted above, especially region-specific effects of armed conflict, that disproportionately affect grade 3 (3) girls in the intervention group.

Proceeding from macro-level explanations to the micro-level, it is worth considering why progress is attenuated more severely in literacy than in numeracy among grade 3 (3), intervention-group learners. One possible explanation is that attenuated learning in numeracy is a result of *distressed* girls not performing their best, while attenuated learning in literacy is a result of not only distress, but also *forgetting* – i.e. the fact that language skills are very easily lost if they are not practiced frequently.

To unpack this explanation further, there are two primary mechanisms hypothesised by which girls' scores might remain the same or even decrease over time (despite girls staying in school): distress and forgetting. Distress describes a condition that could be caused by many different stressful life-circumstances (e.g. traumatic experiences related to armed conflict; the death of a parent, caretaker, or friend) that might put a given learner under sufficient stress that they would perform poorly in class and on an assessment, *even when it comes to skills that they have acquired and mastered*. The implied counterfactual is that the removal of this stressor would allow a given student's performance to quickly return to 'normal' pre-distress levels of performance, because they have the requisite skills. Forgetting describes the loss of a skill or skills due to insufficient practice or utilization of those skills. The implied counterfactual is that new learning (re-learning) and practice will be necessary in order for the learner to return to 'normal' levels of performance.

These two mechanisms of distress and forgetting can be at work individually, but they are also likely to be at work simultaneously for some of the most at-risk learners, especially those learners who have experienced conflict-related stress and trauma. The very conditions that produce distress (e.g. the death of a family member) are also likely to produce forgetting (e.g. by girls attending school less or having less time to study, or losing a family member who used to help them with their homework). Following from this observation, the hypothesised reason for differential rates of negative progress between literacy and numeracy is that forgetting tends to happen faster and more severely with language-related skills such as reading, whereas the loss of math skills is less likely and is expected to happen at a much slower rate.

Readers will recall from the discussion above that a much larger proportion of the girls in the intervention grade 3 (3) sample may have been affected by conflict and resulting school closures. This finding does not help to explain why girls in the comparison group have performed so much better than average, but it does help to explain the extreme under-performance of grade 3 (3) girls in the intervention group in terms of why they would be subject to a higher degree of distress and forgetting than girls in the comparison group.

#### *Testing the effect of individual interventions on learning*

This sub-section presents an analysis of two, more specific project interventions that were not uniform across the intervention group and that thus permit a more detailed analysis of what worked and what did not work from a programming perspective. This analysis is motivated by the absence of a clear intervention-effect in the difference-in-differences analysis presented earlier, as well as the significant regional and other subgroup-wise differences noted above. For example, indicators of low teaching quality (based on caretaker reporting in the household survey) are significant predictors of learning outcomes in the barriers analysis above. This finding suggests that teaching quality is still an important barrier to be addressed, but it does not reveal whether ongoing attempts to train teachers are having their intended effect.



Because numeracy boost training occurred as part of the initial SOMGEP-T intervention and with a specific set of teachers who teach specific grade-levels, it is possible to isolate the students in the sample who were in the relevant grades at the schools where trained teachers teach and who thus had the highest likelihood of benefitting from trained teachers.<sup>93</sup> The intervention group itself is thus subdivided into girls who benefitted from trained teachers and those who did not. Because it was determined that there was no detectable intervention-effect in the difference in differences analysis above, the analysis of teacher training can treat teacher training as though it were the only intervention that occurred. This simplifies the analysis and allows for difference-in-differences estimation that is identical to the analysis performed earlier.

The hypothesis being tested is whether girls belonging to the classes of teachers who received numeracy boost training had numeracy scores that improved to a greater extent (from baseline to midline) than the improvement among students who were in the classes of teachers who did not receive the training. The result of the difference in differences test suggests that the students of numeracy-boost trained teachers did improve over and above those with teachers who did not receive the training, but not to a statistically significant degree. On the basis of this finding, it is not possible to conclude that the numeracy boost training is having the intended effect, but it is certainly possible that the full positive effect of the training is not yet observable because of the relatively short duration of the project up to this point as well as the fact that, for the benefits of training to be detectable at the level of student learning, the benefits of training need to first accrue to teachers (as they practice new skills) and then be transmitted to students during their ongoing learning.

The other specific intervention that can be tested separately involves Girls' Empowerment Forum (GEF) participation, because GEF participation is unevenly distributed among intervention girls. A specific challenge when testing the hypothesis that GEF participation positively affected learning outcomes is that GEFs existed at the time of the baseline, and so exposure to GEF participation cannot be treated in the same way as exposure to teaching from literacy-boost trained teachers.

Thus, the simplest hypothesis that can be tested for GEF participants is whether girls who reported GEF participation at the midline would have learning scores that were higher than all other girls (who did not participate in GEF) at the midline. The result of this hypothesis test is that, within the midline sample, girls who reported attending GEF had significantly higher numeracy and literacy scores than those who said they did not attend GEF. The robustness of this result was also checked on the baseline and midline samples by interacting reported GEF participation (in either wave of the study) with baseline versus midline. The result of this test converges with the results of the simpler test, showing that girls who reported attending GEF at the midline had scores that were significantly higher than those who did not attend GEF at the midline, as well as being higher than all baseline girls (both those who attended GEF and those who did not).<sup>94</sup>

---

<sup>93</sup> It is understood that the benefits of teacher training are meant to diffuse from trained teachers to the rest of the teachers at the school. However, the short duration of project programming thus far suggests that there has not been sufficient time for benefits of training to diffuse, and thus the effects of numeracy boost training are likely to be localized within the grade-levels and classes taught by the trained teacher.

<sup>94</sup> For the results of these specific statistical tests, please see the statistical appendix.



### 5.3. Testing the TOC – intermediate outcomes and learning outcomes

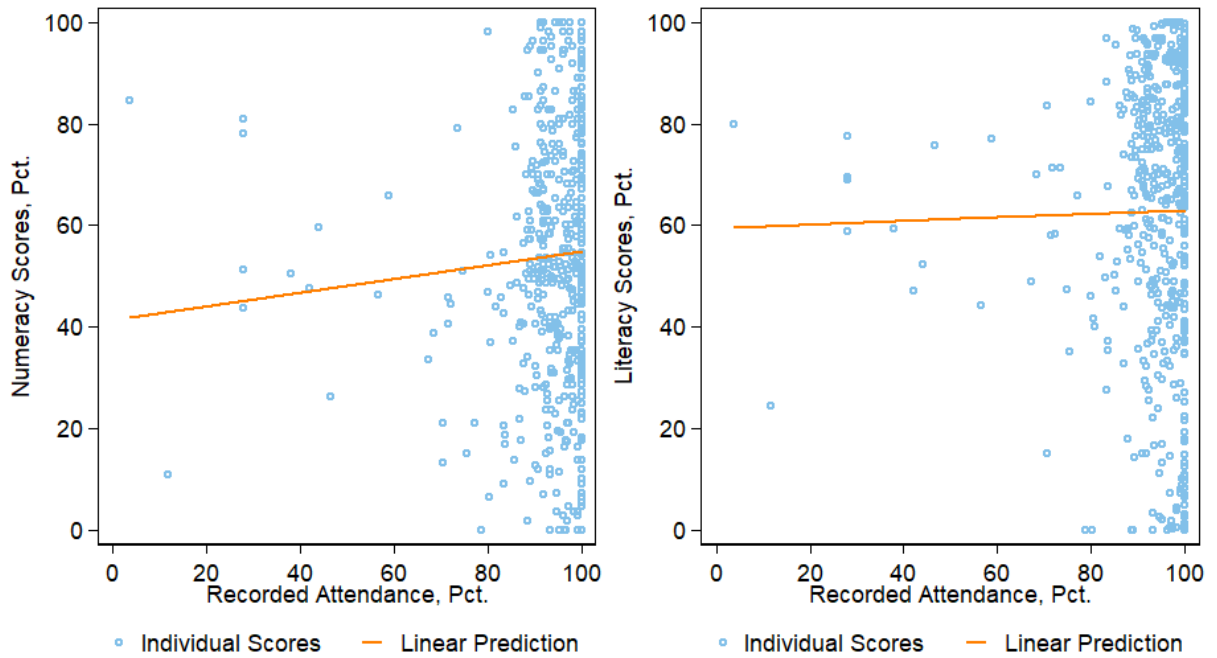
This section explicitly tests key causal linkages between intermediate outcomes and learning outcomes. The project TOC implies that each of the intermediate outcomes (to be covered in detail in section 5 below) have the potential to contribute to learning outcomes. Each of these implied contributions can be tested quantitatively through an analysis of the level of correlation between intermediate outcomes and learning outcomes and cross-checked through triangulation with the available qualitative data. Below, we examine attendance and teaching quality, which are the two intermediate outcomes that have the most direct link to learning outcomes.

#### *IO 1: Attendance Rates and Girls' Learning*

The ToC hypothesizes that higher attendance rates will in turn lead to better learning outcomes. When a regression is run of learning outcomes, that is numeracy and literacy scores, on attendance variables in the household survey as well as in the school survey, we find that attendance is consistently positively correlated with both literacy and numeracy outcomes. We triangulate our findings on the basis of two independent measures of attendance: from school records (gathered as part of the school survey) and from caregiver's estimates.

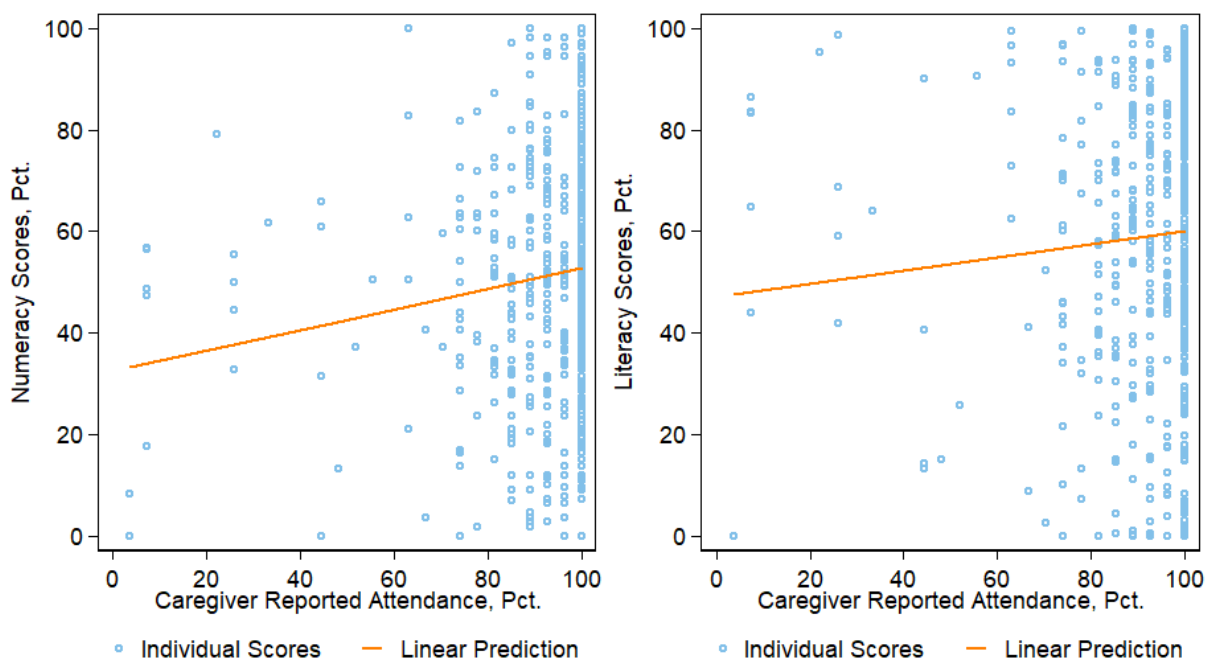
The panel of graphs below visualises the results of regressions of school-record attendance against learning scores. For both numeracy and literacy, the slope of the regression line is positive, although the correlations are not statistically significant.

Figure 9: School Record Attendance as a Predictor of Learning



In order to have a means of cross-checking attendance records and triangulating data, caregivers were asked to estimate the proportion of days during the past month when girls missed school. The panel of graphs below visualises the relationship between attendance as estimated by caregivers and learning outcomes. As with the analysis of school records above, the slope of the regression line is positive indicating a positive relationship between attendance and learning, and the correlation is statistically significant for numeracy (at  $p = 0.005$ ).

Figure 10: Caregiver Estimated Attendance as a Predictor of Learning



The findings above are convergent and provide significant support to the project theory of change by confirming the linkage between attendance and learning outcomes, and also suggesting that project interventions geared toward increasing attendance are likely to have a corresponding, positive influence on learning outcomes.

### IO 3: Teaching quality and girls' learning

The ToC hypothesises that improved skill-specific teaching quality – i.e. addressing specific teaching skill gaps that are reflected in girls' learning – will translate into improved learning outcomes for girls. The barriers analysis above provides a basic test of the linkage between teaching quality and learning outcomes. In the analysis above, low teaching quality and low-quality school leadership (as reported by the primary caregiver) are both statistically significant predictors lower literacy and numeracy outcomes. In addition, reported use of corporal punishment in the classroom (which is also a specific behavioural indicator of poor teaching quality) is associated with lower than average numeracy and literacy outcomes (to a statistically significant degree for literacy). As a whole, these findings are consistent with baseline findings, and these findings re-emphasise the fundamental link between teaching quality and learning that is foundational to the TOC.

When considering interventions to improve skill-specific teaching quality, the above analysis of the effects of numeracy boost training suggests that girls who received the training are performing better than average, but their scores have not yet improved to a degree that the quantitative results are statistically significant. Assuming that the improvements from numeracy boost training are still in the process of being absorbed by trainees and diffused to girls (and potentially to other teachers who did not receive the training), our hypothesis is that numeracy boost training, if truly effective, will be a stronger and statistically significant predictor of numeracy outcomes by the time of the second midline.

## 5.4. Transition Outcome

The second core outcome of all GEC-T projects is transition, defined as the successful continuation of girls in school or suitable alternatives. This section presents the overall findings for transition obtained at the midline, focusing on two distinct areas of analysis. First, we present descriptive findings regarding transition rates and pathways at the midline; these findings are descriptive insofar as they do not make claims about the project’s impact, but simply describe the current state of transition in ways that can guide adjustments to programming and project design. Much of this analysis focuses on specific transition pathways, rather than net transition rates, as pathway-specific findings can be used to target specific girls or adjust programming in response. Second, we provide findings regarding the project’s impact on transition thus far, employing a difference-in-differences methodology that accounts for the nature of sampling at the baseline and midline to provide the most rigorous possible analysis under the circumstances.

To briefly summarize the performance of the project with respect to transition, the table below reports topline transition results relative to the targeted improvement since baseline. Based on transition rates observed at the baseline, the target improvement in transition rates at midline was 8 percentage points. Note that this target applies to children of all grades or ages, in line with guidance from the FM. As we describe in more detail below, the project did not meet its transition target.

*Table 37: Transition performance against targets*

Transition Target (over and above comparison group)	Project Impact	Pct. of Target Achieved
8%	-1.8%	-22.5%

Before discussing key transition results, it is important to review the nature of sample selection at the midline and how this influences the transition rates we calculate. As discussed previously in Section 2.3, the sample drawn at baseline was a random sample of girls, aged 10-19 years, in sampled communities. These girls constitute the “cohort girls” to be tracked over the life of the project. Rather than follow up with all girls in the cohort, the midline evaluation targeted the subset of girls who were in school at the baseline, electing not to re-contact OOS girls at this stage.<sup>95</sup> The consequences of this decision are essential to understanding the analysis we undertake:

<sup>95</sup> As noted elsewhere, another shift occurred in the sample of schools themselves: at midline, a slightly smaller set of schools was chosen from among the set of schools sampled at baseline (to be specific the baseline sample had 76 schools, compared to 63 at the midline). This change impacts nearly all of the analysis in the report, which employs a comparable sample by, effectively, ignoring the 13 additional schools when studying change over time.

- The midline and baseline samples, without adjustment, are fundamentally incomparable, as the midline excludes girls who had dropped out of school or never enrolled in the first place. These girls have the lowest transition rate the baseline – as re-enrolment in school or alternative education is not very common – producing lower net transition rates in the baseline.
- At the midline, we cannot measure the full spectrum of transition pathways, which include OOS girls re-enrolling, transitioning into employment or alternative education, or staying out-of-school. Our analysis of transition is, therefore, truncated, as we illustrate in more detail below.

To provide a like-for-like comparison between baseline and midline, we take a subset of the baseline sample that is limited similarly to the midline. Specifically, we study the subset of girls at the baseline who were in school *in the year prior to the baseline* (i.e. at time  $t-1$  if the baseline is time  $t$ ). Our midline sample is equivalent: a set of girls who were in school in the year prior to the midline (i.e. at baseline). To fix ideas, let the baseline be time  $t$ ; girls are included in our baseline sample if they were enrolled at time  $t-1$ ; girls are included in our midline sample (conducted at time  $t+1$ ) if they were enrolled at time  $t$ .<sup>96</sup> It is important to note that this no longer constitutes a panel design – while many of the same girls are in both subsamples, there is not complete overlap.<sup>97</sup>

There are further subtleties that should be noted and suggest mild caution in interpreting our results. A small subset of “types” of girls – where type is a specific transition pathway over a two-year period – is included in our adjusted baseline sample but not our, otherwise equivalent, midline sample. The table below decomposes the samples into different broad transition pathways, indicating pathways would result in inclusion in the adjusted baseline and midline samples, respectively. The difference that remains between the two samples concerns girls who lived outside of sampled communities in the year prior to the baseline or midline surveys, but are enrolled in sampled communities at the time of the evaluation. Inclusion in the baseline sample was based on presence in the community *at the baseline*, meaning that girls who were enrolled at time  $t-1$  in another community can be included in the sample, as long as they were present at the baseline.<sup>98</sup> In contrast, inclusion in the midline sample is not based on exclusively on presence in the community at the midline, but also on presence in the community at the baseline, because the sample is drawn from cohort girls interviewed at baseline. For this reason, girls who were enrolled in time  $t$  (baseline) in another community and migrate to a project community at time  $t+1$  (midline) are excluded, as they could not be part of the cohort sample established at time  $t$ .

Pathway of Girl	Included in Adjusted Baseline	Included in Adjusted Midline
<b>Stable Residents</b>		
In-school last year; OOS this year	Yes	Yes
In-school last year; in-school this year	Yes	Yes
OOS last year; OOS this year	No	No
OOS last year; in-school this year	No	No

The school sample does not have any special consequences for the calculation of transition rates, beyond those that apply to other outcomes in this report.

<sup>96</sup> Enrolment at time  $t-1$ , in the case of the baseline sample, is measured retrospectively by asking girls and their caregivers whether the girl was enrolled the year prior to the baseline.

<sup>97</sup> To provide just one example, a girl who was enrolled at time  $t-1$ , the year prior to the baseline, but out-of-school at the baseline would be included in the baseline subsample but not the midline subsample.

<sup>98</sup> Girls who were out-of-school in another community at time  $t-1$  were included in the full baseline sample but are not eligible for inclusion in this adjusted transition sample, which is limited to girls who were in-school at time  $t-1$ .

<b>In-Migrants</b>		
In-school last year in different community; OOS this year in this community	Yes	No
In-school last year in different community; in-school this year in this community	Yes	No
OOS last year in different community; OOS this year in this community	No	No
OOS last year in different community; in-school this year in this community	No	No
<b>Out-Migrants</b>		
In-school last year in this community; OOS this year in different community	No	No
In-school last year in this community; in-school this year in different community	No	No
OOS last year in this community; OOS this year in different community	No	No
OOS last year in this community; in-school this year in different community	No	No

Finally, note that the threat to inference outlined above is largely obviated by the difference-in-differences design. Subtle changes to the precise nature of the sample between baseline and midline may result in biased estimates of the over-time change in transition rates. However, any differences in the sample from baseline to midline affect intervention and comparison areas equally. If intervention and comparison areas are equally impacted, the difference-in-differences estimates of project impact on transition rates are not biased by this change. In the absence of differential levels of in-migration to either intervention or comparison communities, the problems outlined above should not affect the difference-in-differences estimates provided below. To summarize: aggregate changes in transition over time should be interpreted very cautiously; differenced changes, in contrast, should not be affected and can be interpreted as estimates of project impact.

### Transition Pathways

SOMGEP-T considers the full spectrum of transition pathways, often differentiating between highly-specific outcomes, such as gainful and non-gainful employment. The transition pathways studied by the project are reproduced in the table below from the baseline evaluation. To reiterate our discussion above, the pathways which cannot be measured in this evaluation are emphasized in red text. The remaining transition pathways were measured at the midline and are included in our comparison of baseline to midline transition rates. In short, transition can only be measured among girls who were in-school in the year prior to the evaluation point.

*Table 38: Transition Pathways Matrix*

	<b>Baseline point</b>	<b>Successful Transition</b>	<b>Unsuccessful Transition</b>
	Enrolled in Grade 3, 4	<ul style="list-style-type: none"> <li>In-school progression</li> </ul>	<ul style="list-style-type: none"> <li>Drops out of school</li> </ul>

<b>Lower primary school</b>		<ul style="list-style-type: none"> <li>• Drops out but is enrolled into alternative learning program</li> </ul>	<ul style="list-style-type: none"> <li>• Remains in same grade</li> </ul>
<b>Upper primary</b>	Enrolled in Grade 5, 6, 7, 8	<ul style="list-style-type: none"> <li>• In-school progression</li> <li>• Moves into secondary school</li> <li>• Moves into ALP</li> </ul>	<ul style="list-style-type: none"> <li>• Drops out of school</li> <li>• Moves into work, but is below legal age</li> </ul>
<b>Secondary school</b>	Enrolled in Grade F1, F2, F3, F4	<ul style="list-style-type: none"> <li>• In-school progression</li> <li>• Enrols into technical &amp; vocational education &amp; training (TVET)</li> <li>• Gainful employment</li> <li>• Moves into ALP</li> </ul>	<ul style="list-style-type: none"> <li>• Drops out of school</li> <li>• Moves into employment, but is unpaid or otherwise exploited</li> </ul>
<b>Out of school</b>	Dropped out	<ul style="list-style-type: none"> <li>• Re-enrol in appropriate grade level in basic education</li> <li>• Enrol in alternative learning program</li> <li>• Engages in wage/self-employment</li> </ul>	<ul style="list-style-type: none"> <li>• Remains out of school</li> </ul>

The overall transition rate among the midline sample was 81.5 percent. Successful transition was defined, in line with the pathways described above, as in-school progression from grade to grade, enrolment in alternative education, enrolment in vocational training, or transition into gainful employment for those girls of appropriate working age. In practice, only three outcomes were common in the midline sample: in-school progression, repeating a grade, and dropping out of school entirely. Only two girls in the midline sample enrolled in alternative education during the prior year, and no girls enrolled in vocational training or entered employment, according to the criteria defined above. The total midline sample size for assessing transition rates was 648 girls, as girls who could not be re-contacted but were replaced were not included in the sample.<sup>99</sup>

The tables below provide a breakdown of transition pathways by age and intervention status. The first table reports transition pathways for girls in intervention community at the midline; the second table reports the same information for comparison communities. These tables highlight several key findings: as noted above, very few girls transitioned into non-school alternatives, such as employment or alternative education. This is at least partially a function of the fact that the girls in question were enrolled at the

<sup>99</sup> Replacement girls were chosen directly from schools, meaning that their transition rate is, by definition, 100 percent.



baseline, while OOS girls may be more likely to seek employment or vocational training than girls who have recently dropped out. In addition, the biggest difference between intervention and comparison transition patterns is that girls are less likely to drop out in intervention communities, but are more likely to repeat a grade. While repeating a grade does not “count” as successful transition, keeping girls who would otherwise drop out in school – even if they are repeating a grade – is a positive outcome, and suggests some progress as a result of the project.

Finally, transition rates at the midline do not have a monotonic relationship with age – rather, successful transition increases from age 11 to 13, remains highest in ages 13 and 14, and then declines among older girls. Finally, dropout rates tend to increase with age: across intervention and comparison areas, midline dropout rates among girls 11-13 were just 3.0 percent, compared to 15.6 percent among girls 16 and above.

Table 39: Transition rates and pathways, intervention group

Intervention group (girls)										
		Transition Pathway								Transition rates
		Successful Transitions				Unsuccessful Transitions				
Age at Midline	Sample size (#)	In-school progression	Enrolled in TVET course	Enrolled in informal education	Gainful employment (of age)	Remains in same grade	Underage employment	Non-Gainful employment	Drops out of school	Successful transition rate per age (%)
11	60	86.7	0	0	0	11.7	0	0	1.7	86.7
12	55	78.2	0	0	0	18.2	0	0	3.6	78.2
13	58	89.7	0	0	0	10.3	0	0	0	89.7
14	63	87.3	0	0	0	9.5	0	0	3.2	87.3
15	42	71.4	0	2.4	0	23.8	0	0	2.4	73.8
16	32	81.3	0	0	0	9.4	0	0	9.4	81.3
17	21	81.0	0	0	0	14.3	0	0	4.8	81.0
18	6	50.0	0	0	0	16.7	0	0	33.3	50.0
19	4	50.0	0	0	0	0	0	0	50.0	50.0
20	2	100.0	0	0	0	0	0	0	0	100.0
<b>Overall</b>	<b>343</b>	82.2	0	0.3	0	13.4	0	0	4.1	82.5

Table 40: Transition rates and pathways, comparison group

Comparison group (girls)										
		Transition Pathway								Transition rates
		Successful Transitions				Unsuccessful Transitions				
Age at Midline	Sample size (#)	In-school progression	Enrolled in TVET course	Enrolled in informal education	Gainful employment (of age)	Remains in same grade	Underage employment	Non-Gainful employment	Drops out of school	Successful transition rate per age (%)
11	53	73.6	0	0	0	18.9	0	0	7.5	73.6
12	48	81.3	0	0	0	14.6	0	0	4.2	81.3
13	59	86.4	0	1.7	0	10.2	0	0	1.7	88.1
14	45	86.7	0	0	0	4.4	0	0	8.9	86.7
15	43	81.4	0	0	0	11.6	0	0	7.0	81.4
16	37	73.0	0	0	0	10.8	0	0	16.2	73.0
17	10	70.0	0	0	0	10.0	0	0	20.0	70.0
18	7	71.4	0	0	0	0	0	0	28.6	71.4
19	2	50.0	0	0	0	0	0	0	50.0	50.0
20	1	100.0	0	0	0	0	0	0	0	100.0
<b>Overall</b>	<b>305</b>	80.0	0	0.3	0	11.5	0	0	8.2	80.3

The distinction in common transition pathways across age groups is demonstrated more clearly in the figure below, which plots the frequency of three common transition pathways among a combined baseline and midline sample of girls. A combined sample is useful for illustrating these trends due to its larger size – enrolment in alternative education is sufficiently uncommon that a larger sample is needed to detect differences across age groups. In this sample – as in the tables above, which focus on only the midline data – transition rates are correlated with age, but do not decrease monotonically. Instead, transition rates increase through age 14 and then decline continuously as age increases.

The figure below makes clear a more interesting set of relationships between transition and age. First, the share of girls who move from school into alternative education only loosely increases with age and remains low across the board.<sup>100</sup> In the qualitative interviews, there is some evidence that girls and women are pursuing alternative learning paths, but these appear to mostly be girls or women who previously dropped out due to marriage, pregnancy, or to help their families. In comparison, both the number of girls who drop out and repeat grades is strongly related to age, but the rate of dropouts and repeated grades are inversely related to one another. As girls get older, they are much more likely to drop out of school, but are less likely to repeat grades. One likely explanation of this finding is that girls who are struggling in school at younger ages tend to repeat grades but stay enrolled, while older girls who struggle are more likely to simply drop out, rather than repeat grades.

There is some evidence in the qualitative interviews that older girls who are in lower grades face stigmatization, which could provide one potential explanation for why girls prefer to drop out rather than repeat grades. When interview participants were presented with a hypothetical situation in which an older girl is attending classes and asked to share what they feel the outcome would be for the girl, one respondent shared that she feels the girl would face ridicule from other students: "...it can happen that her classmates will pass some words such as 'You are older than us and we are in the same grade,' and also elder people can say to her 'You are older than other children, and you are still in grade one.'"<sup>101</sup> This attitude does, however, appear to be changing, as a number of female respondents also shared that they, themselves have returned to school to receive an education, although it is not clear in all cases whether these respondents are referring to alternative learning programs or to the formal schooling system. One mother explains how she left school and did not enrol again until after having children: "I was studying in this school, then in grade 6 I left the school because my mother needed my help and to take care of our livestock. Then I got married and I had children, and I always wanted to continue my education. Now I joined the school again, and always when I'm in school I think about my children. I remember when I was young before I got married – there were no worries and always education was in my mind."<sup>102</sup>

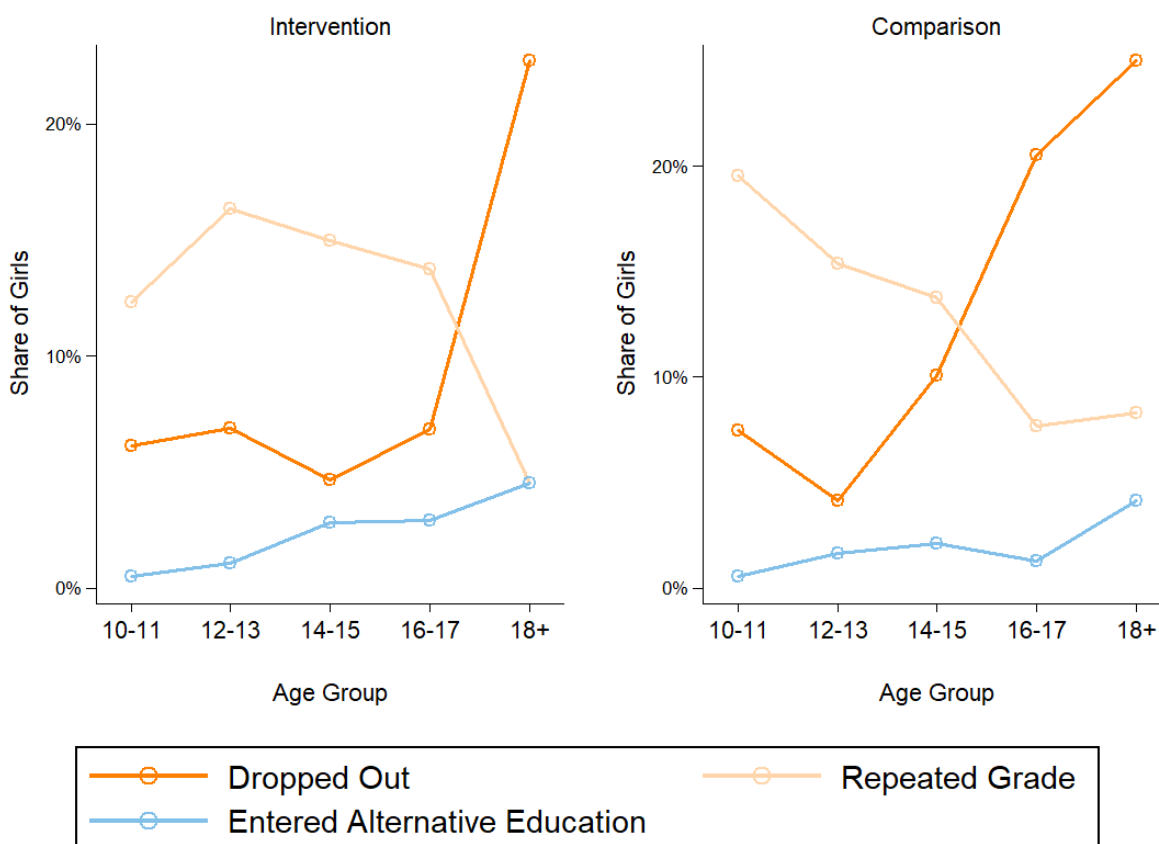
---

<sup>100</sup> Note, however, that this graph includes only girls who were in-school in the prior year; to the extent that OOS girls are most likely to move into alternative education, this graph understates the share of girls who fall into this pathway.

<sup>101</sup> FGD – Storytelling

<sup>102</sup> FGD – Storytelling

Figure 11: Transition pathways as a function of age, by intervention status



Compared to the relationship between age and dropout rates, grade is a less meaningful predictor of dropout rates in our sample, likely because age and grade are only loosely related in the project’s context. Indeed, the data document a number of girls 15-17 years old who dropped out in the past year from grades 3-5 (4-6), which is significantly behind the expected grade level for their age.

Nonetheless, other data suggest that, while dropouts occur at every grade level, they are arguably most frequent – in terms of sheer volume, at least – in the lower grade levels. The figure below utilizes data from classroom headcounts to document the precipitous decline in the number of girls enrolled as grade level increases. In grade 1 in the midline, the typical class has 20.4 girls enrolled; enrolment drops to 14.5 girls in grade 2, 11.7 girls in grade 3, and continues declining to just 7.9 girls in grade 8.<sup>103</sup> An important conclusion from this analysis is that older girls in early grades (i.e. those who have fallen behind) are at the highest risk of dropping out of school. By falling behind, girls may be confirming family members’ idea

<sup>103</sup> A similar trend is observed in boys’ enrolment, though boys’ enrolment is marginally higher than girls’ at every grade level. Note that this analysis is not significantly affected by the local population of girls at different grade levels, because there is little reason to think that communities should have systematically more young girls than older girls. While secondary schools may draw from a different (larger) population than primary schools, the grades documented here are included in the same school and therefore draw from the same underlying population of available students.

that investing in their education is not worthwhile; further, by repeating grades, the total cost of education through some arbitrary level (e.g., grade 8) is increased, as school fees must be paid for the same grade twice. Repeating a grade or otherwise falling behind in school, therefore, appears to be a good indicator of being at-risk for dropout.

Again, and as noted above, the qualitative data supports this explanation. When asked which types of community members face the greatest barriers to enrolling in school, one REO answered: “Girls that are getting a bit older in age and do not have any educational experience. For example, there was a girl that I met who said that she is becoming too old and finding it difficult to sit with grade 3 students who are young children.”<sup>104</sup> Additionally, we know from the qualitative results that parents face a tradeoff in sending their girls to school – without their young girls staying home to help with chores, mothers must take on the household chore burden. For mothers who could otherwise be engaging in income-generating activities, this tradeoff is particularly unattractive. One mother explains the considerations that mothers face in sending their girls to school: “Yes, the lack of finances, busy mothers who work in the market, and girls having to stay home to do household chores are contributing factors.”<sup>105</sup> As was also noted in the learning outcomes section, mothers therefore are particularly keen to see evidence that their children are benefitting from receiving an education in some tangible way, and if they do not see this evidence, they may not see the value in continuing to enrol their girls in school. One REO notes that “I used to have parents say to us that the livestock will run away if their children do not supervise them, as they did not understand the importance of the education. Parents also use to say look at the university graduates that do not have any jobs.”<sup>106</sup> One mother specifically stated that she pulled her daughter out of school when she realized she was not learning well.

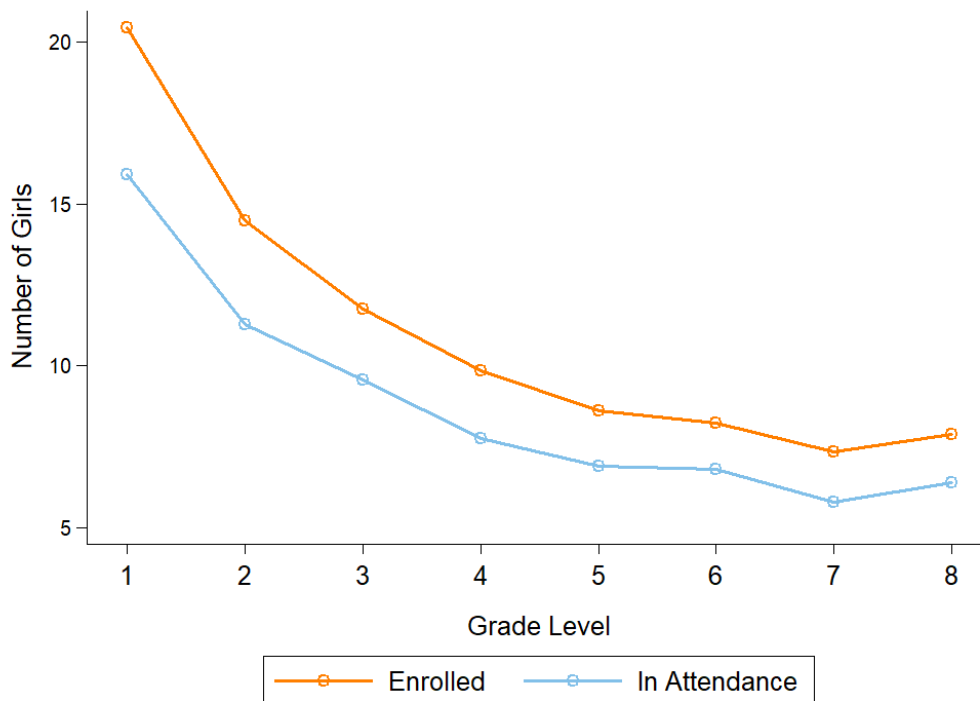
---

<sup>104</sup> KII – REOs

<sup>105</sup> FGD – Mothers

<sup>106</sup> KII – REOs

Figure 12: Count of enrolled girls per classroom, by grade level



### Overall Transition

The results presented above distinguish between a number of distinct transition pathways. However, as a close review of the transition pathways tables reveals, only three pathways were common in the data: in-school progression, repeating a grade, and dropping out. This is not surprising, given that the sample for assessing transition in this analysis is limited to girls who were in-school in the year prior, meaning that re-enrolment is not possible and other pathways – enrolment in alternative education or joining the labour force – are less likely. In this section, we focus on successful transition as a binary outcome, rather than distinct pathways; the following section disaggregates transition according to a variety of population subgroups.

Our primary estimate of project impact at the midline is the change in transition rates between baseline and midline in intervention communities relative to comparison communities. The table below reports the transition rates at baseline (left) and midline (right), disaggregated by intervention status, and provides the difference-in-differences estimate of project impact as well. As the table shows, transition rates increased in both intervention and comparison communities from baseline to midline. This is an important result, and it is not driven by major changes in the sample composition between the two periods.<sup>107</sup> Transition rates in intervention communities increased by 5.9 percentage points; in comparison

<sup>107</sup> As discussed previously, the baseline and midline samples analysed here are constructed to be as comparable as possible.



communities, transition rates increased by 7.7 percentage points. Combined, the estimated effect of the project, at midline, on transition rates is -1.8 percentage points.<sup>108</sup>

*Table 41: Overall transition rates and difference-in-differences*

Group name	Intervention transition rate (Baseline)	Control transition rate (Baseline)	Intervention transition rate (Midline)	Control transition rate (Midline)	Difference-in-Differences	Transition Target	Pct. of Target Achieved
In-school girls	76.6	72.6	82.5	80.3	-1.8	8.0	-22.5%

The magnitude of the increase in transition rates in comparison communities is noteworthy but unexplained – it may stem from structural changes, such as an easing of drought conditions or an improvement in the local security situation. It is also possible that the outsized improvement in the comparison group is driven by reversion to the mean: the comparison group lagged the intervention group by four percentage points at the baseline. It is possible that the baseline performance of the comparison group represented a break from its underlying trend line, and the rapid improvement from baseline to midline was an artifact of reverting to that trend line. Mean reversion is a common threat to valid causal inferences in a difference-in-differences framework, particularly when intervention and control groups are not selected with random assignment and in cases where it is not possible to verify the evolution of pre-intervention trends in the two groups.<sup>109</sup> Unfortunately, the data do not allow us to test whether mean reversion has influenced the results. It is also worth noting that the substance of these results may change in future evaluation waves, when all cohort girls from the baseline are re-contacted.

The results in the table above are derived from a simplified difference-in-differences approach, which is akin to a linear regression model without control variables. Based on evidence from the baseline evaluation and from results – discussed in detail below – analysing population subgroups at the midline, it is clear that transition rates differ systematically by age, geographic region, and as a function of other girl- and school-level characteristics. As a robustness check, we estimated difference-in-differences within a linear regression framework, incorporating a diverse set of control variables in different models. Specifically, our analysis incorporated binary indicator variables for each age level, each zone, disability status (girls with any of the measured disabilities), female-headed households, girls who feel unsafe on their way to school and/or at school, girls who have a relatively higher – spending more than a half day, typically – chore burden at home, and girls who are single or double orphans. Our estimate of impact on transition rates does not change meaningfully in any of these specifications. In the simplified estimate provided in the table above, the estimated impact of -1.8 percentage points is not statistically significant at any conventional level ( $p = 0.79$ ). In the more expansive models, impact estimates ranged from -1.7 points to -3.6 points, but every estimate was indistinguishable from zero or a null effect. We leave

<sup>108</sup> The results presented here are the unweighted means for each subgroup (e.g., the intervention group at baseline). We elected not to use weights in this analysis because it is unclear what the goal of weighting would be; more specifically, it is not clear of what population weights would make the sample representative. With that said, we also performed this analysis using weights that ensure each school is weighted equally in each of the baseline and midline samples – such that a girl in a school with five respondents would be weighted twice as heavily as a girl in a school with ten respondents – and the results we substantively unchanged.

<sup>109</sup> Winship, Christopher, and Stephen L. Morgan. 1999. “The Estimation of Causal Effects from Observational Data.” *Annual Review of Sociology* 25: 694.

discussion of the correlation between these control variables and transition rates aside and return to it in the subgroup analysis in the next section.

### *Heterogeneous Transition Impacts*

Beyond the aggregate effect of the project at the midline, we also investigated its impact within distinct subgroups. Note that this analysis is separate from the subgroups analysis below, which studies the relationship between a girl’s characteristics, for instance, and the likelihood of transition. Rather, we study whether the project’s impact on transition rates differs across subgroups; in other words, we perform the difference-in-differences analysis among subgroups and report the subgroup-specific impact estimate. The idea is that aggregate effects could mask meaningful variation between, for instance, zones – variation that could show a null effect in the aggregate is driven primarily by one zone, or one age group.

The table below reports baseline and midline transition rates and difference-in-difference estimates, disaggregated by age group. Note that age groups are defined by a girl’s age at the baseline. Comparing baseline and midline rates in the intervention group, it appears that the gains in transition rates documented previously are concentrated in the 12-13 year old age group; among the 14-15 year old cohort, intervention group transition rates fell by 1.9 points, and fell more dramatically – by 7.0 points – among girls in the 16+ year cohort. In control communities, no equivalent pattern emerged: transition rates increased most from baseline to midline among girls in the middle two age groups, but large gains were still observed among girls aged 16 years and older.

*Table 42: Transition rates and difference-in-differences estimates by age group*

Age Group (age at baseline)	Intervention transition rate (Baseline)	Control transition rate (Baseline)	Intervention transition rate (Midline)	Control transition rate (Midline)	Difference-in-Differences	Overall Sample Size
10-11 years	79.3%	72.7%	82.6%	77.2%	-1.2%	472
12-13 years	71.6%	76.9%	88.4%	87.5%	6.2%	521
14-15 years	78.9%	69.3%	77.0%	77.5%	-10.1%	364
16+ years	79.7%	66.7%	72.7%	70.0%	-10.3%	157

Differences between the intervention and comparison groups in terms of where gains were seen from baseline to midline underpin the difference-in-differences estimates in column 6. While the overall difference-in-differences estimate for transition was -1.8 points, the project showed relative improvements in transition among the 10-13 year old cohort (i.e. girls aged 10-13 years at baseline), and the opposite among girls 14 and up.<sup>110</sup>

<sup>110</sup> Unsurprisingly, none of the age group regressions produced statistically significant results, because the sample sizes were comparatively smaller (ranging from 226 to 516 observations) once the sample was split.

### *Intervention-Specific Impacts on Transition*

Throughout our discussion of the project's impact on transition rates, we have focused on estimates of overall project impact, rather than the impact of specific interventions. On one hand, this approach is logical, because our primary interest is in the impact of the “basket of interventions” that a single project applies to a community or school. This is especially important if, as in the case of GEC-T, we expect the impact of individual interventions to be conditional on one another.<sup>111</sup>

On the other hand, SOMGEP-T's efforts are not monolithic, and it is worth studying which interventions are having the most positive effect. Girls do not receive a uniform set of interventions – indeed, some self-select into portions of the project's activities, such as enrolment in ALPs or participation in the Girls Empowerment Forums – and understanding the relative impact of different interventions or packages of interventions is worthwhile.

Unfortunately, it is often difficult to parse the effects of different interventions, because they typically overlap with one another, and – in many cases – we lack the high-quality data on exposure to different interventions that would be needed. In addition, small sample sizes can become a problem when there are many interventions, because the limited evaluation data is then split into many subgroups of interventions for comparison.

As discussed in the learning section above, exposure to two project activities – GEF participation and enrolment in classrooms where teachers have received numeracy pedagogical training – are well-documented, either via data provided by CARE International or through data collected as part of the evaluation. For the analysis in this section, we focus on GEF participation, because it has the strongest theoretical relationship to transition. Higher-quality mathematics teaching may result in higher transition rates in the long-term if girls and their parents feel they are learning more or that education is a worthwhile investment. But this mechanism is diffuse. In contrast, GEF participation is expected to increase girls' life skills and self-esteem, and may make significant impacts in a shorter timeframe.

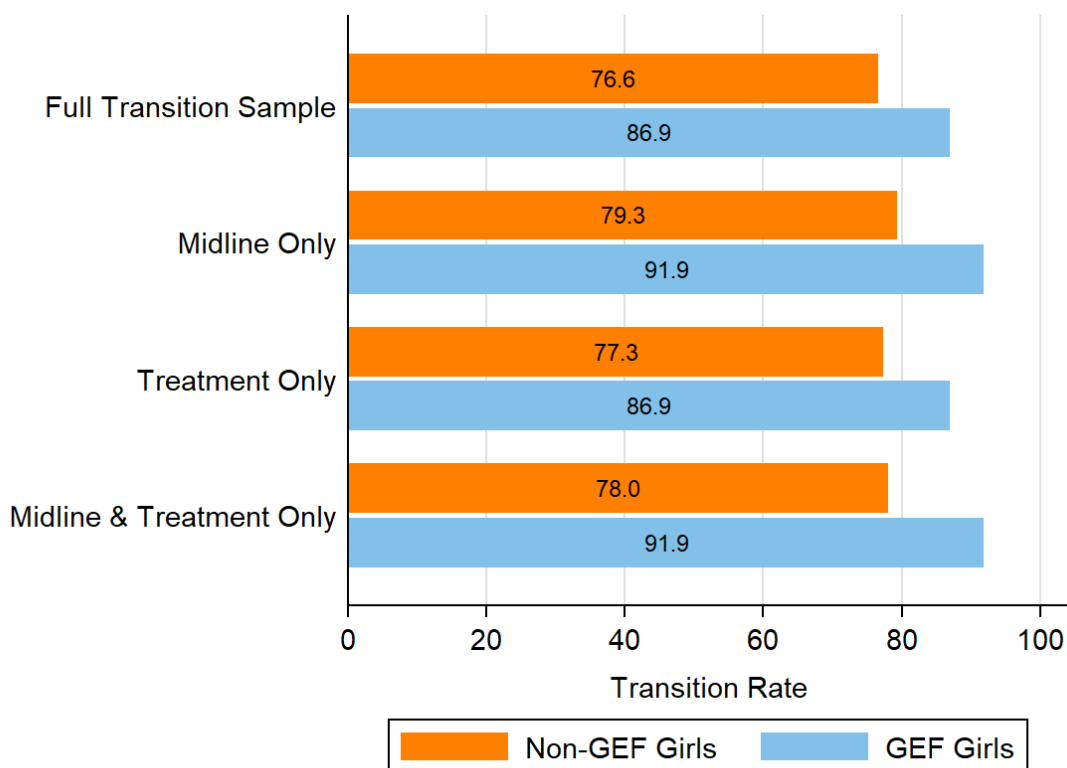
We coded GEF participation on the basis of two questions directed at girls: whether they had heard of the GEF and activities it sponsored, and whether they had ever participated in a GEF activity. Girls in intervention schools who professed to have participated in a GEF activity are the targeted group, who we refer to as GEF girls.

As with learning outcomes, GEF girls outperform other groups in terms of transition rates. When we consider the full transition sample – including baseline and midline evaluations and both intervention and comparison schools – intervention girls are 4.8 percentage points more likely to transition successfully than comparison girls. But when we look only at GEF girls, they are 11.0 points more likely to transition than comparison girls, and 9.6 points more likely to transition than intervention girls who have not participated in the GEF.

---

<sup>111</sup> One such case would be an effort to improve household livelihoods or provide cash assistance, alongside activities meant to raise awareness regarding the importance of girls' education and improve community attitudes toward girls' education. To the extent that we think the positive impact of cash assistance is conditional on positive attitudes among parents, our foremost goal is to estimate the effect of the overall intervention, as conceived of by the project.

Figure 13: Transition rates as a function of GEF participation, in varying samples



In fact, as the figure above illustrates, GEF girls outperform their peers in terms of transition no matter how we parse the sample: compared to all non-GEF girls, they are 10.4 percentage points more likely to transition successfully; compared to all non-GEF girls in the midline, they are 12.6 points more likely to transition; and compared exclusively to other intervention girls in the midline, they are 13.9 points more likely to transition. In every case – and in regression models that incorporate control variables for zone, age, intervention status, round, and cluster standard errors by school – the difference between GEF girls and non-GEF girls is statistically significant at the 5 percent level or better.

While the positive relationship is clear between GEF participation and both learning and transition outcomes, the reason for this relationship is not clear. The hypothesised process is that GEF participation would increase girls’ empowerment or self-esteem (in ways that would potentially be measured through YLI score), which would in turn lead them to participate more in the classroom and improve learning, and also to take more control over everyday decisions that might affect their chances of successful transition. However, as noted above, GEF participation is not predictive of higher YLI scores, and YLI scores are, in turn, not predictive of higher learning or transition rates. Thus, the hypothesised mechanism of GEFs increasing girls’ self-esteem does not appear to be active. In the absence of this mechanism, there is no evidence supporting a clear explanation for the positive relationship between GEF participation and learning and transition outcomes. One possibility that may be worth testing with additional qualitative data in the future is that girls are selecting into GEF participation on the basis of having a more enabling home-environment (a factor that would not necessarily be reflected in individual-level assessments of self-efficacy or YLI score). If this explanation is true, it is this supportive environment at home that is

encouraging them to join GEF, and also that same environment that is enabling them to fare better than average in terms of studying at home, learning at school, and transitioning successfully.<sup>112</sup>

### Target setting for the transition outcome

Given the nature of the transition samples analysed here, we have opted not to set targets on the basis of transition rates observed at this evaluation point. Instead, the targets for the second midline and the endline have been set based on the benchmark transition rates obtained at baseline. There are two reasons for this decision: first, future evaluations are unlikely to limit their sample to in-school girls exclusively, so setting targets based on a sample that includes only in-school girls will lead to incorrect targets and expectations going forward.

Second, using the benchmark transition rates from the baseline provides consistency throughout this report, the baseline evaluation report, and the outcomes spreadsheet. In both the baseline report and the outcomes spreadsheet, transition targets are calculated against the benchmark established at baseline. In addition, our analysis of transition rates earlier in this section judged performance against a target (8 percent) based on that same benchmark. For the sake of consistency and clarity, and because future evaluations will employ a sampling strategy closer to that utilized during the baseline, it is preferable to set targets based on the benchmark transition rates established at the baseline.

The transition targets in the table below are based on the FM's guidance, in line with the discussion above.

*Table 43: Transition targets for future evaluations*

	Evaluation point 3	Evaluation point 4
Target generated by the outcome spreadsheet	10	10

## 5.5. Subgroup Analysis of Transition Rates

This section reports transition rates among various subpopulations of interest, defined by geography, disability, economic status, and other girl- and household-level characteristics. We also report transition rates for girls who face hypothesized barriers to educational attainment, such as discomfort at school, the use of corporal punishment, poor reported teaching quality, and unsafe or long journeys to school.

The goal of this subgroup analysis is two-fold. The first goal is to determine whether particular “types” of girls face especially high hurdles to successful transition. This is a simple test of the Theory of Change, as some of the barriers considered in our analysis are specifically cited by the project’s ToC, especially linking intermediate outcomes to primary outcomes. In addition, this analysis can serve to identify areas in which the project’s implementation could be usefully adjusted by, for instance, emphasizing the teacher training component of the intervention in areas where teaching quality is especially poor and is associated with lower transition rates. The second goal is to assess the extent to which the project has had subgroup-specific impacts. Building on the aggregate results reported above, the subgroup analysis

<sup>112</sup> Girls are not randomly selected into participation in the GEF, and it is possible that the most motivated girls, those with the highest self-esteem, or those with the most supportive families are more likely to join. In short, it is possible that GEF participation and successful transition are jointly determined by a girl’s characteristics or other confounding factors. Regardless of the precise mechanism, the findings in this section should motivate additional focus on GEFs as mechanisms for improving girls’ education, and for additional research into their effects.

reports difference-in-differences models limited to subsamples (where sample sizes are sufficiently large to permit this analysis). By assessing project intervention effects within specific subpopulations, we shed light on whether the project benefits only those girls who face the fewest barriers, whether project impacts are spread evenly across the sample, or whether the project benefits disadvantaged girls most.

The initial subgroup results are provided in the table below. Membership in each subgroup – the first column of the table – is binary, and transition rates are reported for the midline sample as a whole in the second column. The third column reports the change in transition rates for each subgroup from baseline to midline, using a like-for-like comparison between the two.<sup>113</sup> The fourth and fifth columns repeat this analysis, restricted to girls in intervention communities – i.e. the fourth column reports transition rates for subgroup members in intervention schools and the fifth column reports the change in transition rates for subgroup members in intervention schools from baseline to midline. The sixth column provides the difference-in-differences estimate of project impact for this subgroup, with significant results (at the 10 percent level) denoted by an asterisk.<sup>114</sup> Finally, the rightmost column reports the number of girls in the midline transition sample (and the number of intervention girls in the midline sample) who fall into the subgroup. To facilitate comparison of each subgroup to the broader sample, the top row provides the overall transition rate in the midline and midline intervention samples, respectively.

*Table 44: Transition rates at midline, by subgroup*

	Midline Transition Rate	Change in ML Transition Rate from BL	Midline Intervention Group Transition Rate	Change in ML Intervention Group Transition Rate from BL	Estimate of Project Impact (Diff-in-Diff)	Midline Sample Size (Intervention Group Sample Size)
<b>Characteristics</b>						
All girls	81.5	6.8	82.5	5.9	-1.8	648 (343)
Living without both parents	77.2	7.6	68.8	-9.4	-35.2	57 (32)
Mother tongue different from LOI	97.6	1.3	96.3	1.6	1.6	41 (27)
<b>Disability</b>						
Vision impairment	100	25	100	33.3	N/A	3 (1)
Hearing impairment	.	.	.	.	N/A	0 (0)
Mobility impairment	100	40	100	25	N/A	1 (1)
Cognitive impairment	.	.	.	.	N/A	0 (0)
Self-care impairment	.	.	.	.	N/A	0 (0)
Communication impairment	100	50	.	.	N/A	1 (0)

<sup>113</sup> As noted previously, the raw baseline and midline samples differ systematically in ways that would bias our estimation of transition rates. Therefore, we use restricted subsamples that are comparable from both rounds: girls who were enrolled in the year prior (time  $t-1$ ) to data collection. For most details, see the discussion surrounding the aggregate transition results above.

<sup>114</sup> In each case, we specify a linear regression of successful transition on a intervention X round interaction term (the difference-in-differences estimand) and a vector of age group dummy variables, restricting the sample to girls who fall into a given subgroup. Standard errors are clustered by school. Subgroups without results in this column have sample sizes too small to permit reasonable interpretation of the results.

Mental health impairment	70	11.9	72.9	18.8	15.2	90 (48)
Anxious	75	15	75	13.1	-4.8	76 (40)
Depressed	60.3	7.2	63.6	17	19.8	58 (33)
Any disability	71	8	73.5	11	6.4	93 (49)
<b>HOH and Caregiver Characteristics</b>						
HOH no wage-earning occupation	83.5	9.3	82.8	6.7	-5.5	334 (180)
HOH no education	77.9	1.4	80.6	1.9	0.63	335 (175)
HOH female	77.5	0.7	77.6	-1.7	-5.1	293 (156)
HOH Pastoralist	66	-5.6	57.9	-10.1	-4.1	47 (19)
Carer no education	79.9	4.5	82.4	6.1	3.0	358 (176)
<b>Household Assets</b>						
Owens camels	82.4	3.8	88.9	12	14.6	68 (27)
Owens medium-sized livestock	81.7	7.5	83.4	6.6	-1.9	438 (217)
Owens small livestock	87.5	9.7	88.9	13.9	N/A	32 (9)
Owens mobile phone	82.3	7.6	83.8	6.8	-1.7	560 (309)
Access to water reservoir/storage	80.7	7.8	84.1	9.6	5.0	389 (233)
Owens land	81.2	7.4	81.7	6.4	-2.0	474 (240)
<b>Poverty</b>						
House is informal/temporary structure	76.7	-4.3	76.2	-0.7	5.3	43 (21)
Gone to sleep hungry most days	73.1	-5.9	60	-20	N/A	26 (10)
Gone without enough clean water most days	78.9	-5	74.2	-6.2	-1.9	57 (31)
Gone without medicines or medical intervention most days	81.1	5.2	78	1.3	-7.9	122 (59)
Gone without cash income most days	81.1	5.1	78.4	-0.8	-13.7	122 (74)
<b>Migration</b>						
Displaced or moved in past 12 months	100	40	100	0	N/A	2 (2)
Household migrates seasonally	71.4	-4.2	33.3	-39.4	N/A	7 (3)
<b>Other</b>						
High chore burden (whole day spent on chores)	69.8	13.5	71.9	23.5	21.8	63 (32)
Married	20	-71.7	0	-88.9	N/A	10 (4)



Mother, under 16	66.7	-33.3	.	.	N/A	3 (0)
Speaks LOI poorly	91.3	-3	100	14.3	N/A	23 (7)

As with learning outcomes, it is not possible to draw conclusions about the impact of specific types of disability on transition rates, because most of the disability type subgroups include too few girls for effective analysis. For instance, just three girls in the midline transition sample have impaired vision – drawing inferences regarding transition rates from such a small sample is not recommended. The exception concerns mental health: 90 girls in the midline transition sample were classified as being either depressed or anxious – or both – and we find lower transition rates among girls who fit either description. While girls in this category still have lower transition rates, their performance has risen from the baseline: in a comparable sample of baseline girls, transition rates among girls with mental health disabilities were just 58.1 percent, compared to 70.0 percent at midline. This shift is driven almost entirely by improvements among the intervention group, where transition rates have increased by 18.8 percentage points, compared to a 3.5 point increase in comparison schools. The difference-in-differences result bears this out, though the small sample size means that even a large project impact in this subsample is not statistically significant.

Girls with both physical disabilities and mental health concerns were captured in the qualitative interviews, providing us with some idea of the barriers these girls might face to continued enrolment. The results suggest that disabled girls do not attend school as regularly as their peers due to their disabilities, and there is even some evidence that disabled girls may miss large chunks of a given school year if they are feeling particularly unwell or receiving treatment. As one girl explains, “...there was a time when I was feeling pain in my leg and I did not attend school for about one month.”<sup>115</sup> Girls with disability are therefore forced to rely on their classmates for assistance after absences. Frequent absences, combined with the general difficulties girls with disability face in focusing and learning while in school, likely put girls with disability at greater risk of falling behind than their peers. Additionally, although almost all the girls with disability that were interviewed appear to receive significant support from their peers and reported that they do not feel excluded, there is evidence that girls with disability face ridicule from their peers. One REO explains, “Having a disability does have an effect on the girl’s attendance at school. She will be subject to bullying or abuse from other students, especially from boys, which could cause her to drop out of school.”<sup>116</sup> Of particular concern is the fact that a number of girls with disability appear to have not been born with their disabilities but rather to have been victims of violence. One girl reports that a boy damaged her eye, and another that a crazy man injured one of her hands so much that she can no longer write.

The characteristics of a girl’s head of household and caregiver are less predictive of transition rates than her own disability status. The results here are mixed: girls in households where the head of household does not earn a wage are actually slightly more likely to stay in school and progress from grade to grade, while those in female-headed households and those in which the head of household has no formal education are marginally less likely. Where a girl’s caregiver lacks education, there is a small drop in transition rates as well. Overall, transition rates for these subgroups have increased from baseline, but there is no evidence that this stems specifically from project impact, because the increases are spread approximately evenly between intervention and comparison communities.

---

<sup>115</sup> KII – GWD

<sup>116</sup> KII – REO

The role of household economic status is similarly mixed, based on the results presented in the table. Household asset ownership is not correlated with systematically higher transition rates – girls in households that own at least one camel or cow, or own at least one goat or sheep, or own a mobile phone are – at best – only very slightly more likely to remain in school and progress through the grades than girls in households that lack these assets. Where household economic status does seem to matter is when it takes the form of acute economic deprivation: among households that frequently experience hunger (go to bed hungry at most days), transition rates are just 73.1 percent. Indeed, even if we take a broader view of food insecurity, households that have gone to bed hungry even one or two days over the past year have lower transition rates (73.8 percent) than those who never experienced food insecurity of this kind (85.2 percent).<sup>117</sup>

In practice, these results likely understate the importance of economic conditions in shaping transition rates. Economic deprivation is multifaceted, and utilizing one measure at a time does not take into account the way in which deprivation on one metric can interact with deprivation on another. For instance, going with a cash wage is problematic, but it is especially difficult for households that do not have a sizable herd of livestock. In addition, many standard metrics of poverty – such as the construction materials of one’s home – are long-term in nature, and do not reflect short-term economic fluctuations in response to sudden job loss, drought, unexpected medical bills, and so forth.

We expect that short-term economic hardship should have an outsized impact on transition rates, because it is more closely tied to the contemporaneous decision of whether to enrol one’s child in school in a given school year. To establish a multidimensional measure of short-term economic deprivation, we combined data on livestock ownership, and the frequency with which households go to bed hungry and/or lack sufficient clean water for home use.<sup>118</sup> While livestock ownership may seem an odd choice of indicator, we feel it captures an important aspect of short-term economic hardship, because animals are readily sold as a buffer during hardship, and smaller herds are therefore correlated with greater short-term deprivation. This contrasts with indicators such as household construction quality, which are not easily drawn down. Our scale ran from a minimum of zero to a maximum of nine points, with higher scores indicating greater deprivation.

For the purpose of this analysis, we classified households as economically deprived if they scored above five points on our index, which covered 16.1 percent of households in the midline sample. Among households defined as deprived, transition rates at midline were 72.1 percent, while rates among non-deprived households were 83.2 percent, a statistically and substantively significant difference ( $p = .008$ ). This difference is sensitive to the somewhat arbitrary choice of five points as our cutoff, but the broad conclusion remains the same: a lower threshold for deprivation (above four) produces a gap between the deprived and non-deprived of 4.3 points, indicating that deprivation is still associated with lower transition rates. These results were supported by the qualitative findings and consistent with findings regarding

---

<sup>117</sup> This broader definition is useful, because it expands the sample size of affected households in the midline sample from just 26 to 187.

<sup>118</sup> Specifically, we scored households based on their ownership of large (camels or cattle), medium (goats or sheep) and small (chickens, etc.) livestock. Households with no large animals received 2 points; households with 1-2 large livestock received 1 point; households with 3 or more large livestock received 0 points. A similar score was established for medium livestock. For smaller animals, where large numbers are rarer, households were given 1 point if they did not own any small livestock and 0 points otherwise. With respect to hunger: households that experienced hunger many or most days over the past year received 2 points; households that experienced hunger one or two days received 1 point and households that never experienced hunger received 0 points. Finally, households that had insufficient clean water many or most days received 1 point; all others received 0 points.

learning outcomes. For instance, inability to afford school fees, whether tuition fees or fees associated with paying for uniforms, exams, or materials, was mentioned most frequently by respondents as a barrier to enrolment or continued enrolment. And, as with learning outcomes, transition rates are lower among pastoralist households and those experiencing economic deprivation.

The analysis of the deprivation index also highlights the fact that pastoralist households have been disproportionately affected by the drought and its continued effects, in line with what one would expect theoretically. While the extent of deprivation has fallen in non-pastoralist households from baseline to midline, pastoralist households are *worse off* in November 2018 than they were one year prior.<sup>119</sup> The gap in experiences between pastoralist and non-pastoralist households from baseline to midline is also reflected in transition outcomes: at baseline, there was just a 3.4 percentage point difference in transition rates between these groups, with girls in non-pastoralist households slightly more likely to remain enrolled or transition successfully. At the midline, this gap had widened considerably: from baseline to midline, transition rates in non-pastoralist households rose from 75.0 to 82.7 percent; in pastoralist households, transition rates fell from 71.6 to 66.0 percent, producing a gap at midline of 16.7 percentage points.

In most of the remaining subgroups included in the table above, the sample sizes are too small to draw firm conclusions about their impact on transition rates. One exception is the role of a girl's chore burden – the midline sample included 63 girls who were viewed as having a particularly onerous set of household chores, which took up an entire day.<sup>120</sup> Consistent with the baseline findings, girls in this group experienced lower overall transition rates, though they also showed significant improvement over the baseline, rising from 56.4 percent in the baseline transition sample to 69.8 percent in the equivalent midline sample. Gains were particularly large in intervention, relative to comparison, communities (23.5 versus 3.2 points), though the small sample size in the intervention versus comparison group suggests these results should be interpreted cautiously.<sup>121</sup> Chore burden was still mentioned by midline qualitative interview participants as a major barrier to attendance. As one teacher noted, “In order for the girls to continue their education, the parents must reduce the house chores that they send to the girls. They make the breakfast in the morning, they cook dinner at night, and they may not get the chance to read, so

---

<sup>119</sup> This finding makes clear that the drought has differential impacts by geography and household type, with rural, pastoralist households most affected by the drought and having the most difficulty recovering from its impact. While urban households are also affected, their impacts are generally indirect – the loss of income among the extended family, bringing extended family members into their households, or the more diffuse generalized economic impacts, such as higher food prices and slower economic growth, felt across the entire region.

<sup>120</sup> While the sample size is small, the finding regarding girls who speak the language of instruction (LOI) at their school poorly may appear noteworthy, given that girls in this group are much more likely to remain in school. However, in the Somali context, girls who do not speak the LOI well all attend schools in which English or Arabic is the language of instruction (because Somali is universally spoken, in contrast to many other GEC-T contexts). Therefore, this indicator is primarily measuring the type of school a girl attends, rather than any actual linguistic disadvantage vis-à-vis other girls (e.g., a Kenyan girl who speaks Kiswahili poorly but must use Kiswahili in school). To the extent that English- and Arabic-language schools are unusual – better-resourced or otherwise – it is not surprising that a subset of their students perform better than average with respect to transition.

<sup>121</sup> It is also possible that the project's interventions have reduced the number of girls who are tasked with a full-days burden of chores in the first place. From baseline to midline, the share of girls with full-day burdens dropped 2.3 percentage points in comparison communities and 3.6 points in intervention communities. This change is emblematic of a stronger shift – though still small – in intervention communities, relative to comparison communities, toward lower chore burdens for girls over time, which could be the result of changing community attitudes toward the value of girls' education.

parents must be fair to the girls and boys and give the girls the chance to get an education.”<sup>122</sup> The household chore burden can result in tardiness, absence, or lower learning outcomes, as girls have less time to focus on their homework than boys.

The next set of subgroups are defined not by household characteristics or even enduring traits of girls, but by school- or teacher-level characteristics and, in some cases, the manner in which girls interact with those characteristics. The structure of the table is identical to that above, and we report overall midline transition rates in the first row to facilitate comparisons to midline transition rates among specific subgroups that follow. As with the previous findings, caution is warranted – due to the small sample sizes involved – in the interpretation of some of the results. For this reason, we report the number of girls in the subgroup, and the number of intervention girls in the subgroup, in the right-most column, to make clear the strength of the evidence base for particular results.<sup>123</sup>

*Table 45: Transition rates at midline, disaggregated by barriers to transition*

	Midline Transition Rate	Change in ML Transition Rate from BL	Midline Intervention Group Transition Rate	Change in ML Intervention Group Transition Rate from BL	Estimate of Project Impact (Diff-in-Diff)	Midline Sample Size (Intervention Group Sample Size)
<b>Barriers:</b>						
All in-school girls	81.5	6.8	82.5	5.9	-1.8	648 (343)
<b>School Infrastructure</b>						
Difficult to move around school	87.9	9.2	84	1.2	-14.7	124 (50)
Doesn't use drinking water facilities	90	14	79.4	8.5	-6.9	100 (34)
Doesn't use toilet at school	90.6	8.1	84.5	1	-12.6	138 (58)
Doesn't use areas where children play/socialise	87.5	6.6	81	1	-9.5	248 (105)
<b>School Resources</b>						
No computers at school	86.6	4.7	86.1	3.1	-3.6	568 (309)
School does not have learning materials	84	-3.5	83.1	-3.8	-0.5	144 (65)
Not enough seats for children at school	79.3	-2.4	86	3.5	11.9	87 (43)
<b>Teaching Quality</b>						
Disagrees teachers make them feel welcome	93.1	18.7	92.3	13.4	N/A	29 (13)

<sup>122</sup> FGD – Teachers

<sup>123</sup> While there are many competing rules of thumb regarding the necessary sample size for valid statistical analysis, our view is that results from subgroups with fewer than 50 total members should be treated with caution and results from subgroups with fewer than 30 members should be disregarded as unreliable in most cases.

Agrees that they are afraid of teacher	89.5	9.6	89	8.8	-1.7	476 (246)
Agrees teachers treat boys and girls differently in the classroom	88.2	9.5	89.1	7.7	-3.9	262 (137)
Agrees teacher is often absent from class	84.6	9.2	86.8	9.4	-0.4	130 (53)
Teacher punishes students who get things wrong	87.3	6.4	86.5	3.8	-5.7	497 (260)
Teacher uses corporal punishment	89	3.7	88.5	1.6	-4.8	182 (96)
Carer says principal performance is poor	80	2	84.6	10.9	N/A	15 (13)
Carer says teaching at school is poor	66.7	-11.5	80	-11.7	N/A	12 (5)
<b>Gender Equity</b>						
Teacher targets questions by gender	82.1	-6.7	73.3	-18.3	-24.5	56 (30)
Teacher targets difficulty of questions by gender	83.3	-4.8	78.7	-8.8	-10.2	78 (47)
<b>Other Barriers</b>						
Agrees she has no choice in schooling decisions	88.3	7.9	88.2	6	-4.1	540 (279)
Over 30-minute travel time to school	82.9	0.4	76.5	-2.1	N/A	35 (17)
Feels unsafe on way to school	94.7	22.2	100	30.8	N/A	19 (12)
Feels unsafe at school	100	36.8	100	52.6	N/A	11 (7)

The first set of barriers to transition that we investigate concern the school's infrastructure and resources. Importantly, the measures of school infrastructure analysed here are not, strictly, measures of quality or availability; rather, they are self-reports of whether a girl uses facilities – such as a toilet, playground or available water supply – at their school. These indicators, therefore, jointly measure whether a school has a toilet, whether the girl is able to access it, and whether she finds it acceptable for use, in addition to other factors that might inhibit her use of the toilet.

In general, girls who report difficulties with their school's facilities have *higher* transition rates than other girls. Girls who report that they have difficulty moving around the school, and girls who report that they do not use the toilet at their school have transition rates of 87.9 and 90.6 percent, respectively, compared to the sample average of 81.5 percent. In each of these four infrastructure-focused measures, transition rates have increased from the baseline, but the size of this increase has been larger in comparison than intervention communities. A lack of school resources – such as desks and computers – also do not impact transition rates in the way we would expect. Girls in schools that lack these materials are no less likely – and possibly more likely – to remain in school and progress to the next grade.

Teaching quality and teacher demeanour provide another set of counter-intuitive results. At the baseline, girls who reported that their teachers made them feel unwelcome, who reported that they were afraid of their teacher, and who stated that their teacher was often absent all had lower transition rates than average. These relationships, and most other measures that concern teaching quality, are reversed at the midline: in all three of the subgroups noted, girls have higher transition rates at midline than average.<sup>124</sup> Among girls who say that their teacher is often absent, transition rates were 75.4 percent in the baseline sample, but rose to 84.6 percent at midline. The one subgroup in this set that conforms closely to our expectations are girls whose caregivers state that teaching quality at their child's school is poor, but the sample size is too small to draw firm conclusions about the role teaching quality plays in shaping transition rates. As was noted above, there was one mother in the qualitative interviews who specifically stated that she pulled her daughter out of school due to low learning outcomes, which does provide some evidence for this conclusion: "For example, when my daughter was studying there, I found out after 2 years that she did not understand or learn much. I followed up on it and found out that she was not being taught anything at school. So I took her out of school and nobody came to me to follow up on why I had taken her out of school."<sup>125</sup>

Girls' self-perceived empowerment also does not appear to increase transition rates. We identify girls who feel they have no choice in their schooling decisions, who agree a little or a lot with the statement "I cannot choose whether to attend or stay in school. I just have to accept what happens." Among girls who feel they have no control over these decisions, transition rates were 80.5 percent in the baseline sample and 88.3 percent in the midline sample. At both baseline and midline, these girls transition at higher rates than other girls. Given the nature of the question, and the possibility that girls misunderstood it, we also considered alternative measures of perceived empowerment by girls and found similar results: girls who felt they had less control over schooling decisions actually transitioned at higher rates.<sup>126</sup> Another potential explanation is that parents' attitudes toward education are changing as they are exposed to programming. There is certainly evidence of this attitude shift in the qualitative interviews. One mother explains, "I stopped making her help around the house with household chores, so she can prepare her lessons and go to school early, and I support her in any way I can further her education."<sup>127</sup> In the baseline, the responses of some mothers suggested that girls keep their relationships secret and sometimes decide to drop out of school to get married against the advice of adult figures in their lives. Although the results could be due to social desirability bias, it is worth noting that almost all of the mothers interviewed for the midline expressed an interest in keeping their girls enrolled in school. Many

---

<sup>124</sup> These same girls who report negative characteristics of their teachers – such as an unwelcoming nature – score no differently, on average, on learning assessments at the midline. While less surprising than a comparatively higher transition rate, this finding is similarly counter-intuitive, in that we expect worse learning outcomes to accompany poor teacher demeanour or a less welcoming classroom environment.

<sup>125</sup> FGD – Mothers

<sup>126</sup> The question on which this analysis is based involves a double-negative that may have confused girls. In total, 81.8 percent of respondents at the midline expressed strong agreement with the idea that they are powerless over their schooling decisions. This unusual response pattern – we would expect more girls to express mild agreement or disagreement and, in general, for more girls to believe they have control, even if they do not – combined with the nature of the question, led us to question the quality of the underlying data. To check the robustness of the findings, we used a second question, which asked girls who decides whether they will stay in school or return to school. Girls who reported that their family unilaterally decides (i.e. not jointly with the girl) were classified as having no perceived control over schooling decisions. As with the primary analysis, girls who expressed this view actually had higher transition rates than other girls at both baseline and midline.

<sup>127</sup> FGD – Mothers



noted, as the mother above did, that they have taken on the chore burden themselves in order to keep their girls in school.

The findings to this point rely on bivariate analysis, comparing transition rates in a given subgroup – such as pastoralist or female-headed households – to transition rates in the rest of the sample. The limitation of this approach is that it may produce spurious correlations between a given independent variable and transition rates, primarily in the case where the independent variable is simultaneously correlated with other predictors of transition. To reduce – but certainly not eliminate – the likelihood of such spurious correlation, we estimate a series of linear regression models predicting successful transition.<sup>128</sup>

Consistent with the bivariate subgroup analysis above, the results also show that a range of household- and girl-specific characteristics strongly predict transition rates. Arguably the strongest predictor is pastoralism, which is associated with an 18.5 percentage point decline in transition rates – not far off its apparent effect in the bivariate analysis. Girls with disabilities are also much less likely to remain enrolled in school and advance a grade level, with transition rates 12.2 percentage points lower, all else equal. Female-headed households and households in which *both* the head of household and caregiver lack formal education of any kind are both associated with lower transition rates, though the estimated effects are not statistically distinguishable from a null effect. Finally, our index of household economic deprivation is associated with a small but meaningful reduction in transition rates – if a household moves up two points on a 1-9 scale, their daughter’s likelihood of successful transition drops 4.0 points.<sup>129</sup> In additional models, we controlled for age and grade level at baseline with a series of age- and grade fixed effects, but the results were substantively unchanged.

Beyond the model described in the table, we also estimated a model that captured the effect of a girl’s household chore burden: as with the subgroup analysis, girls with a significant chore burden are much less likely to transition successfully. In the bivariate analysis, girls with a heavy chore burden lagged behind other girls by 11.7 points in terms of transition. Once we control for additional factors, this relationship persists, at a 12.8 points drop in transition rates.<sup>130</sup>

## 5.6. Testing the TOC – intermediate outcomes and transition

This section expands on the more general subgroup analysis presented in the previous section by specifically investigating the links between the intermediate outcomes and transition. As with the learning analysis, our goal is to explicitly test the links hypothesized in the project’s Theory of Change. To do so,

---

<sup>128</sup> Note that standard errors are clustered by school.

<sup>129</sup> We also analysed the correlation between transition rates, on one hand, and the quality of one’s home construction and the frequency with which members of the household went to bed hungry over the last year, on the other. In these models, the results were consistent with those presented in this section: both variables were associated with lower, but not statistically significantly lower, transition rates. By aggregating several measures of deprivation into a single variable, the index captures a slightly broader range of deprivation and also gains statistical power, as indices do when they aggregate multiple variables that tend to move together in the same direction.

<sup>130</sup> We estimated a number of additional models, which captured different aspects of the subgroup analysis above. However, because of the small sample sizes associated with most of the subgroups (e.g., girls with more specific disabilities, girls that are married, households that migrate seasonally), most of the relationships are extremely sensitive to the inclusion or exclusion of other variables. The results presented here are robust to a number of different specification choices, which is why we opted for this fairly parsimonious, but still useful model.



we focus on the two intermediate outcomes most logically linked to successful transition: teaching quality, and girls' life skills.<sup>131</sup>

### IO 3: Teaching quality and transition outcomes

The project's Theory of Change links teaching quality to both learning and transition outcomes, though it's connection to learning is arguably more obvious. Nonetheless, better quality teaching should be associated with higher transition rates for a number of reasons: better teachers instil a stronger interest in learning in their students; they also make clear the value of learning both in terms of intrinsic rewards and extrinsic, material incentives to education; finally, they make learning and being at school more fun and more rewarding. These are just a few ways in which higher-quality teachers can promote improved transition rates.

To study the relationship between teaching quality and transition, we focus on girls in the midline transition sample. Because some of the girls in the sample had dropped out by the time of the midline, they did not answer questions regarding teaching quality, nor did their caregiver. However, this information was captured at the baseline from the caregivers, allowing us to check the relationship between perceived teaching quality at the baseline and transition outcomes at the midline. This is especially appropriate, because transition is not an outcome that occurs *at* the midline, but *between* the baseline and midline, when the decision to enrol or not enrol (or advance a grade) is made. Therefore, a caregiver or girl's opinion from the baseline should influence transition rates over the year that follows, if the hypothesized relationship between teaching quality and transition rates is valid.

The table below reports the results of a series of regression models, building on those in the previous section. In each case, we control for zone, disability status, and binary variables indicating whether a household is headed by a woman, a household is pastoralist, and whether either the head of household or the child's caregiver received any formal education. We also control for household economic deprivation. In each model, we incorporate one metric of teaching quality, and test its relationship to transition, after controlling for these additional factors.

In general, these findings mirror those of the subgroup and barriers analysis. Girls whose caregiver states that the quality of their school principal or their daughter's teacher is poor, transition rates are markedly lower than the average. We combined caregivers who cited a poor quality principal or teacher into a single group, owing to the small number of caregivers who fell into either category individually.

Teaching Quality Indicator	Effect on Transition	P-Value
Caregiver reports poor quality school management or teacher	-12.5 points	0.01*
Teacher encourages participation in class	0.1 points	0.97
Teacher uses corporal punishment	4.7 points	0.25
Student is afraid of their teacher	4.5 points	0.21

---

<sup>131</sup> We do not consider the relationship between attendance and transition here because we consider the two outcomes to be too proximate to one another to be of significant interest. Poor attendance almost certainly is related to lower transition rates, but one could argue that they exist on the same spectrum that ranges from enrolment and full attendance, on one hand, to completely dropping out, on the other.

In contrast, when a girl reports that her teacher uses corporal punishment on students in the class, she is no less likely – and arguably *more* likely – to transition successfully, though the correlation is not statistically significant. Surprisingly, teachers who reportedly encourage participation have no effect on the transition rates of their students. Finally, the counter-intuitive findings from the bivariate analysis, which suggested that girls who felt afraid of their teacher were actually *more* likely to continue in school and transition successfully (89.5 percent successful transition, versus 81.5 percent in the sample as a whole) are partially confirmed by the regression analysis. In the final model reported in the table above, girls who agree that they are afraid of their teacher are 4.5 percentage points more likely to transition, all else equal, though this difference is not statistically significant.

The findings in this section do not support a clear story about the relationship between teaching quality and transition outcomes in girls. On one hand, caregivers who report poor teaching and school management quality are associated with lower transition rates. However, other aspects of teaching quality do not seem to have a systematic effect on transition, and poor teaching quality may actually drive higher transition rates in some specific circumstances.

#### **IO 4: Life skills and transition outcomes**

A key goal of the SOMGEP-T intervention is to improve the financial literacy and leadership skills of girls, as a method of empowering them. Empowerment of this kind is theoretically linked to transition outcomes because girls who exercise more control over decisions that affect their lives are more likely to remain enrolled in school. In addition, girls who obtain leadership skills, participate in their local GEF, and gain increased self-confidence may be more likely to *want* to stay in school, regardless of their actual control over the decision. Our expectation is that measures of life skills and self-empowerment should be correlated with higher transition rates, all else equal.

To assess the relationship between life skills and transition rates, we use two metrics of the intermediate outcome life skills. We employ the Youth Leadership Index (YLI) score, introduced previously in this report and discussed in more detail in Section 6.4. We also employ a binary measure in which girls were asked whether they feel they have any choice in whether they stay in school or not. We code girls who report feeling they have no choice as being relatively disempowered.

<b>Indicator of Life Skills or Self-Esteem</b>	<b>Effect on Transition</b>	<b>P-Value</b>
YLI Score	-0.002 points	0.89
Feels she has no choice in schooling decisions	13.3 points	0.21

As with teaching quality and the subgroup analysis of transition that preceded it, we estimate a linear regression model predicting successful transition, and incorporate control variables for geographic zone and a number of household or demographic characteristics. The results are reported in the table above; for the purposes of focusing attention on the relationships of interest, we do not report regression coefficients for the control variables, but include only the measures of life skills. The first row reports a regression model in which transition was regressed on a number of control variables and the YLI score. As the results make clear, there is no apparent relationship between a girl’s YLI score and her likelihood of remaining in school, in this specification. In the second row, we report a similar model, substituting our binary metric in for YLI score. Here we obtain a counter-intuitive result, in which girls who report feeling

they lack decision-making power are 13.3 percentage points more likely to stay in school, although even this large substantive difference is not statistically significant at conventional levels.

In Section 6.4, below, we report results that show no systematic positive relationship between life skills and learning outcomes. Our results regarding transition are similar – of the two metrics of life skills or empowerment employed, neither predicted higher transition rates. It may be the case that the link between empowerment and transition is too indirect, given the many pressures – financial, social, cultural, and otherwise – facing girls and their families. It may also be the case that the project has not had sufficient time for improvements in empowerment to exercise downstream impacts on broader outcomes such as learning and transition.

## 5.7. Sustainability Outcome

This section presents an analysis of project sustainability at each of three levels: community, school, and system. The results of this analysis are presented in narrative form in the subsections that follow. The triangulated analysis was used to generate a qualitative sustainability score (ranging from 0 to 4) for each of the key sustainability indicators identified in the SOMGEP-T Logframe. Sustainability scores for each indicator can range from 0 to 4, in line with the FM's MEL Guidance:<sup>132</sup>

- 0 – Negligible
- 1 – Latent
- 2 – Emerging
- 3 – Becoming established
- 4 – Established

Scores for indicators in the same level – community, school, or system – are aggregated into a single level score; in turn, those levels are averaged to produce a single sustainability score. These scores, for both baseline and midline, are reported in the scorecard table below.

While additional rationale for each indicator's score are provided in the more detailed narrative sections that follow, it is important to note that the scores are admittedly subjective. We assigned scores based on the totality of information available – in some cases, this was limited to a few qualitative interviews, while in other cases it included quantitative data from multiple respondent groups and a range of qualitative data. As a broad rule of thumb, where noticeable but not dramatic changes have been observed since the baseline, we tended to score indicators as one grade above the baseline. Where two indicators in the same level (community, school, or system) both had marginal improvements, we split the difference, again relying on our best, but subjective, judgment.

*Table 46: Sustainability Scorecard*

	Community	School	System
<b>Indicator 1:</b>	Percentage of CECs actively engaged in mobilizing for girls' education through fundraising for payment of additional teachers' salaries and school supplies  <b>Baseline Target:</b>	Percentage of project target schools adhering to implementation standards for ALP, ESL, Numeracy and Remedial classes.  <b>Baseline Target:</b> N/A	Inclusion of ALP in the national non-formal education frameworks  <b>Baseline Target:</b> N/A

<sup>132</sup> See "GEC-T MEL Guidance Part 2," pages 46-52, distributed May 2017.

	<p>N/A</p> <p><b>Baseline Status:</b> 18.8 % of CECs provide support for teacher salaries (18.8% intervention; 0% comparison).</p> <p>21.7% of parents report CECs provided financial support of some kind (fundraising, infrastructure, buying materials, financial support to students) to schools in the last 12 months (24.6% intervention; 18.4% comparison).</p> <p><b>Midline Target:</b> 35%</p> <p><b>Midline Status:</b> 17.0% of CECs provide support for teacher salaries (19.2% intervention; 14.8% comparison).</p> <p>24.7% of parents report CECs provided financial support of some kind (fundraising, infrastructure, buying materials, financial support to students) to schools in the last 12 months (31.6% intervention; 16.8% comparison).</p> <p>47.6% of CECs, per head teachers, raise funds for school improvements (59.4% intervention; 35.5% comparison).</p>	<p><b>Baseline Status:</b> 43.1% of teachers in intervention schools reported using formative assessments. This assessment is limited to the same sample of schools employed at midline, to facilitate comparisons.</p> <p><b>Midline Target:</b> 30%</p> <p><b>Midline Status:</b> 41% of teachers in intervention schools were able to show records of the use of formative assessments.</p>	<p><b>Baseline Status:</b> Not applicable because activities not established yet</p> <p><b>Midline Target:</b> ASLP documents developed in partnership with MoEs (curriculum validated by MoEs)</p> <p><b>Midline Status:</b> Some mention of ALP in REO interviews. Awareness among Ministry officials is clearly growing with increased discussion of alternative learning opportunities including vocational training.</p>
<p><b>Indicator 2:</b></p>	<p>Percentage of parents in intervention schools indicating that CECs are functional</p> <p><b>Baseline Target:</b> N/A</p> <p><b>Baseline Status:</b> 75.6% of parents in intervention communities report a functional CEC,</p>	<p>Not applicable</p>	<p>No. of MOE departments engaged in support of girls' education from National to regional and district levels.</p> <p><b>Baseline Status:</b> Interviews reveal a systemic lack of funds. Local schools do not have enough money to maintain facilities and pay their staff partly because</p>

	<p>compared to 60.6% of parents in comparison communities. The sample used to assess baseline status was limited to the same set of communities studied at midline, to facilitate valid comparisons between the two rounds.</p> <p><b>Midline Target:</b> 80%</p> <p><b>Midline Status:</b> 81.6% of parents in intervention communities report a functional CEC.</p>		<p>CECs do not have the funds to support their schools. The CECs do not have funds because little to no financial support is given to them by the MoE. Some in-kind support from MOEs was provided. 31% of schools had CEC Management plan, taken as proxy of MOE involvement.</p> <p><b>Midline Target:</b> 20% of the department officers assessed</p> <p><b>Midline Status:</b> Two out of six REOs explicitly confirmed the existence of gender units. Ongoing difficulties recruiting and retaining female teachers and high-quality teachers. Mixed evidence on mainstreaming of improved teaching practices, but these appear to be emergent at this phase. Significant increase in evidence of TVET/NFE units creating options for vocational training. Minimal evidence of MOE efforts to encourage female pre-service graduates.</p>
<b>Baseline Sustainability Score (0-4)</b>	<b>1</b>	<b>1</b>	<b>0.5</b>
<b>Overall Sustainability Score (0-4, average of the three level scores)</b>	<b>0.83</b>		
<b>Midline sustainability Target (0-4)</b>			
<b>Midline score (0-4)</b>	<b>1.5</b>	<b>2</b>	<b>1.5</b>
<b>Overall sustainability Score (0-4, average of the three level scores)</b>	<b>1.7</b>		

## CEC Engagement and Financial Support of Schools

Community Education Committees (CECs) are local structures that oversee the operation of a primary school in a given community. The committees typically consist of the school's head teacher, community leaders, and parents. CECs support the operation of schools in a number of ways, by monitoring student and teacher attendance, promoting enrolment and attendance, raising awareness of the importance of education, liaising with religious leaders and other individuals of influence in the community, and providing material support to the school, among other tasks.

Elsewhere in this report, we highlight the role of CECs in school governance and management, including their efforts to monitor enrolment, attendance, and teacher quality. In this section we focus on the role of CECs in sustaining improvements made to schools through their material – financial and in-kind – support. CECs are expected to raise funds from their own resources, by mobilizing community members to donate, and by seeking support from outside organizations, including NGOs and the government. Money raised, alongside in-kind contributions of time, labour, and materials, can be used for a wide range of purposes, some of which we discuss below. But, regardless of the precise intervention by the CEC, the sustainability of schools following the conclusion of SOMGEP-T requires continued material support from CECs.

Our metrics of financial support by CECs are varied, though they differ somewhat from the baseline evaluation, because midline data collection did not include surveys of teachers. At the baseline, teachers were asked to assess the level of support provided to the school by the CEC and the community more broadly over the past year. At the midline, in the absence of data from teachers, we rely on two alternative data sources: the reports of head teachers, who were asked about CEC support in the payment of teacher salaries, scholarships for students, and material improvements to school facilities; and reports of parents on the role CECs play in fundraising, improving school infrastructure, supporting students financially, and purchasing learning materials.<sup>133</sup>

The first outcome we analyse is the extent to which CECs raised money and paid the salaries of teachers in their schools. Teacher salaries are an area of particular need, because teachers are often paid late; over time, this contributes to discontent among the teaching staff, increasing absenteeism and even prompting teachers to move to schools where financial support is more generous.<sup>134</sup> As one CEC member described it, “The greatest challenge this school faces is the teachers that are brought in from other districts and then they don't receive enough salary. As a result they leave the school.”<sup>135</sup> Teachers confirm salary as a primary concern during qualitative interviews, but this concern is shared by CEC members, who recognize the difficulty that teachers face.

---

<sup>133</sup> The former measure, from head teachers, was only collected at the midline. The latter measure has the advantage that parents in both the baseline and midline were asked to describe the activities of their CECs, providing the opportunity to study improvements in CEC engagement over time.

<sup>134</sup> Qualitative and anecdotal evidence makes clear that timely payment of salaries is a problem. However, neither the baseline nor midline evaluation collected data on whether salaries are consistently paid in full, so we cannot say whether teachers are ever deprived of their salaries. In other research conducted in Somalia, the evaluation team has found that approximately 10 percent of teachers report they were not paid the correct amount of salary in the previous month, which suggests that teacher salaries may be affected by both issues – a lack of timeliness and a failure to pay the full salary. However, to reiterate, no data of this kind was collected at SOMGEP-T schools; our evidence limits us to stating that teacher salaries are often delayed.

<sup>135</sup> FGD – CEC Members

To assess the extent of CEC support for teacher remuneration, head teachers were asked to indicate the share of male and female teacher salaries that CECs paid over the previous year. As shown in the figure below, the overall share of salaries paid by CECs is very low. This fact is not surprising, and is not necessarily an indictment of the CECs, because teacher pay in Somalia is handled in widely varying ways from region to region and school to school, and not all CECs plan to or need to support teacher pay directly.<sup>136</sup> For our purposes, the important finding regarding CEC provision of teacher salary support is that it has increased markedly from baseline to midline. At baseline, just 9.2 percent of head teachers reported that their CEC paid a portion of teacher salaries, and the average share paid of female and male teacher salaries was just 0.3 and 1.7 percent, respectively. At the time of midline data collection, however, 17.0 percent of head teachers reported that the CEC paid a portion of teacher salaries, and the average share of female and male teacher salaries paid rose to 1.1 and 5.5 percent. While the change in share of salaries paid from baseline to midline was not statistically significant, it represents a substantively large increase – in the case of male teacher salaries, more than triple the support was provided at midline as at baseline ( $p = 0.14$ ).

It is unclear why CECs appear to be supporting male teacher salaries more than female teacher salaries, as there was no discussion of this phenomenon in the qualitative data and no evidence that CEC actions or attitudes are gender-absent. In fact, the only time female teachers were mentioned by qualitative interview respondents was when they were highlighting the need for more female teachers to improve girls' experiences at school, recruitment, and retention. In this sense, female teachers appear to be quite highly valued and sought after in some communities. One CEC member explains, "There are big challenges in this school; there are a lot of male teachers and we want female teachers that will improve the quality of girls' education."<sup>137</sup> It is again worth noting that the quantitative finding is based on a small sample of schools where the CEC is providing any level of support for teacher salaries – therefore, taken together, the quantitative and qualitative data suggest that this finding is not necessarily indicative of a bias toward supporting male teachers.

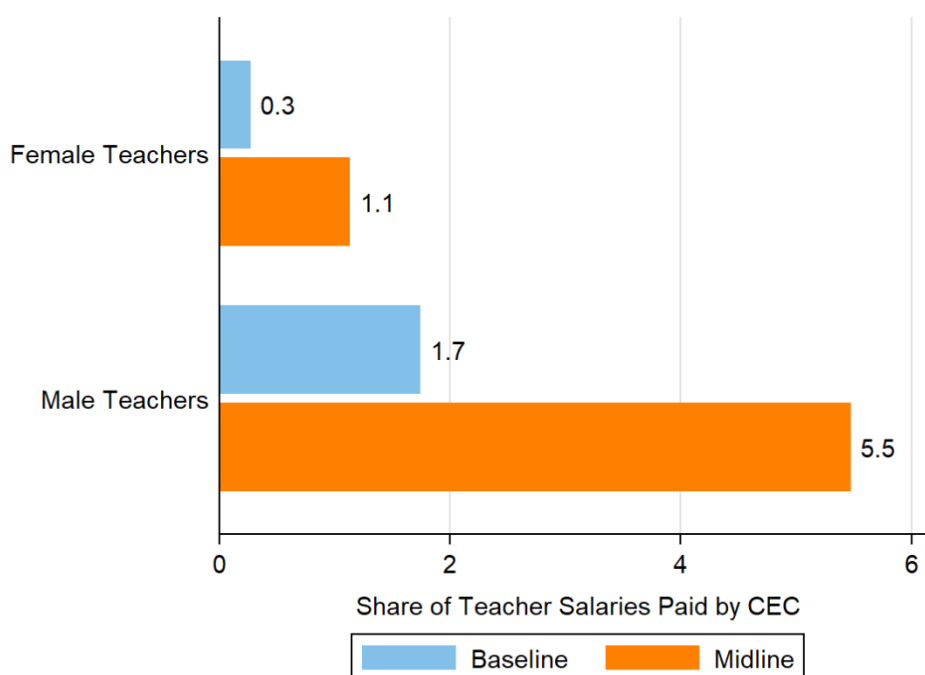
---

<sup>136</sup> In addition to government salaries, some teachers may receive varying levels of support from other sources, ranging from a small top-up incentive to supplement government pay to the payment of their entire salary if they are not a government teacher. Locally-provided incentives or salaries may be funded through the collection of school fees from parents, through CEC fundraising in the community, through donations from local or international organizations, and other sources.

<sup>137</sup> FGD – CEC



Figure 14: Share of teacher salaries paid by CEC, by round



The findings above actually understate the extent of CEC support for teacher salaries, in that they include comparison schools. Given that the project’s sustainability should arguably be focused on intervention schools alone, it is comforting that support for teacher salaries is higher in these communities. At midline, 19.2 percent of head teachers in intervention schools reported that their school received salary support from CECs. The share of support was not dramatically larger than in comparison schools – 5.4 percent of male teacher salaries and 1.5 percent of female teacher salaries – but the set of schools that reported support was broader.

At the same time, it is important to note that the increase from baseline to midline occurred broadly, rather than in intervention schools alone. When we compare trends between intervention and comparison schools, there is no discernible difference in the evolution of salary support provided by the CECs. In fact, if there is a difference in trends, comparison schools – which lagged intervention schools by this metric at the baseline – actually “caught up” to intervention schools somewhat. In our view, the overall rate of salary support is the most important outcome, rather than outsized increases in intervention versus comparison schools, but it is important to highlight the fact that the observed increases may be part of a broader trend toward increased salary support by CECs in the areas where SOMGEP-T is being implemented, as opposed to a phenomenon unique to intervention schools.

Beyond paying a portion of teacher salaries, CECs also take primary responsibility for raising funds for school improvements or repairs. Relative to teacher salaries, CECs were more active in this realm: at midline, 59.4 percent of CECs in intervention areas had raised funds for school improvements of some kind, according to head teacher reports, compared to 35.5 percent of CECs in comparison areas. On the other hand, CECs were somewhat less active in the provision of scholarships to female students. While 58.7 percent of head teachers reported that one or more girls in their school was receiving scholarship support at the midline, the vast majority of this support came from sources other than the CEC. Just 7 of

63 CECs (11.1 percent) at the midline provided scholarship support to female students, and just four of those CECs were in intervention areas.<sup>138</sup>

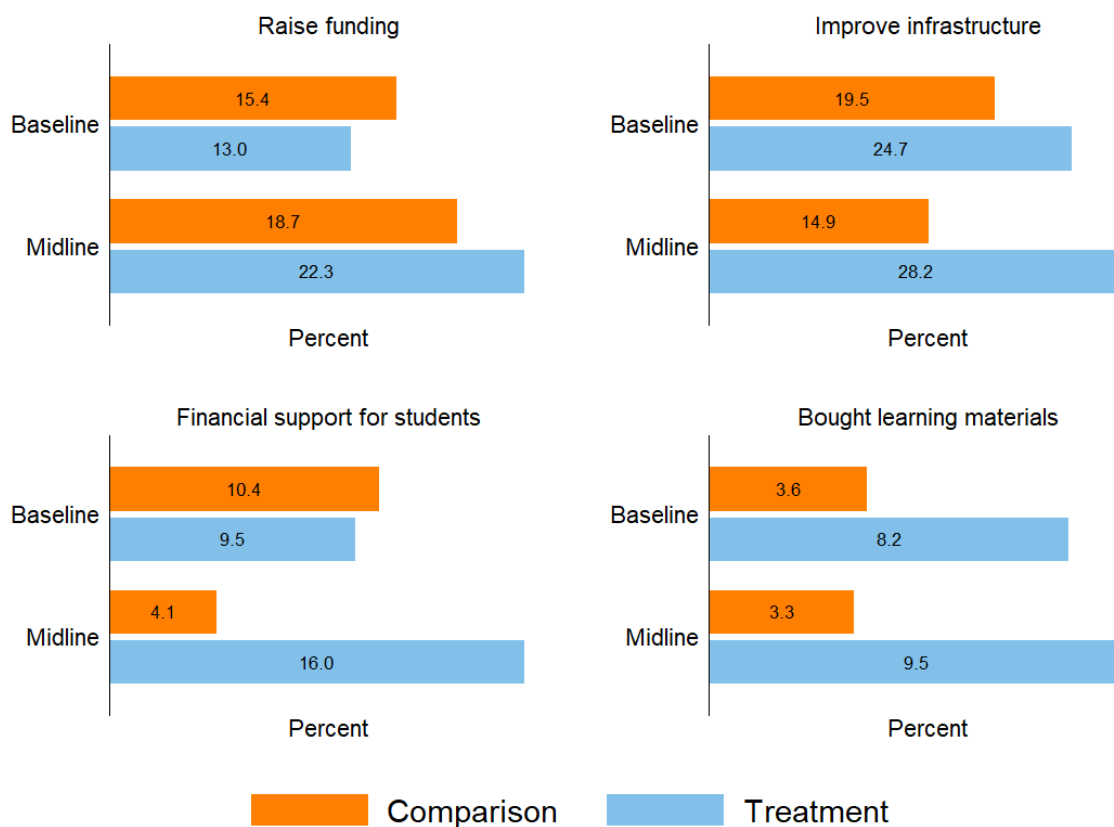
These findings are further supported, and in some cases clarified, by the reports of parents. As shown in the figure below, 22.3 percent of parents in intervention areas said that their CEC had raised funds for the local school in the past 12 months. In general, we should probably expect parents to perceive less action by CECs than head teachers, because parental reports rely on awareness among parents, which is not universal. Parents may not be entirely aware of what is happening at the local school, and may be less aware of who is behind infrastructure improvements, the payment of salaries, etc. Head teachers, in contrast, should be very aware, due to their role within the school and their frequent participation on the CEC itself. According to parents in intervention schools at the midline, CECs are about equally active in raising money for the school (top-left panel), improving school infrastructure (top-right panel), and providing financial support to students (bottom-left panel). The latter finding stands in some degree of contrast to the reports of head teachers: as noted above, head teachers reported the provision of few scholarships to female students; but 18.1 percent of parents in intervention schools at the midline reported that the CEC had provided financial support to students in the last 12 months.<sup>139</sup>

---

<sup>138</sup> Unfortunately, head teachers were not asked to assess CEC support for school improvements or their provision of scholarships at the baseline, so it is not possible to determine whether CEC support in these areas has increased since the project's inception.

<sup>139</sup> Importantly, the gap between head teacher and parental perceptions could stem from differences in the questions posed. Head teachers were asked about female students, while the question directed at parents was gender-neutral. Head teachers were also asked specifically about the provision of scholarships, while parents were asked about "financial support." Financial support could be a conceptually wider term, encompassing the provision of uniforms or other necessarily supplies, while scholarships might imply support specifically in the form of paying a student's school fees. In addition, financial support might be seen to include ad hoc and one-time support, while scholarships could be viewed as more institutionalized and formal support, perhaps on a continuing basis.

Figure 15: Financial support of schools by CECs over the last year, according to parents



The most compelling conclusion that emerges from the figure above is the extent to which financial engagement of CECs in intervention areas has increased over time. In all four areas of financial support (we include in-kind support, such as the purchase of learning materials, because the relevant distinction is between CEC actions that require financial resources and those, such as awareness-raising or attendance monitoring, that do not) CECs in intervention areas have improved from baseline to midline. This trend is especially noteworthy, because the opposite trend is observed in comparison areas – on three of the four metrics, comparison-area CECs have *reduced* their financial engagement, according to parents, and comparison-area CECs’ performance on the other metric (raising funds) has changed minimally from baseline to midline. In contrast, the share of parents in intervention areas reporting that their CEC provides financial support to students has increased from 9.5 percent to 16.0 percent over the last year, and parental reports show substantively meaningful positive progress on all four measures.<sup>140</sup>

<sup>140</sup> This analysis uses data collected from parents of cohort girls; parents of ALP girls are excluded, because no comparable comparison communities were sampled. Note also that the analysis is limited to those parents who indicated that their community has a functional CEC (i.e. answered in the affirmative when asked whether their school has “a CEC that helps with school-related matters”). When we include parents who denied that there is a functional CEC – coding them as indicating that their CEC *did not* provide any of these four types of financial support to the school in the last 12 months – the results are substantively unchanged.

Relative to the results thus far, the qualitative data paint a similar, but somewhat more nuanced, picture of CEC engagement in terms of material support. While 19.2 percent of intervention schools at midline reported salary support from their CEC, CEC members themselves reported a more extensive level of support. Most of the CECs that participated in FGDs indicated that they contributed to teacher salaries; in two cases, they reported providing approximately 1,000 USD and 23,000 USD for teacher salaries.<sup>141</sup> Other CECs, though presumably less prolific, also indicated that they paid portions of teacher salaries from money they raised.<sup>142</sup> In line with the quantitative results, though, a number of CEC members emphasized their support for school improvements and repairs, even if these repairs often strained the CEC's finances and posed risks of not being completed. As one CEC member described, "The school requires repairs. There is one room in the school where we began building but couldn't afford to finish because we ran out of money."<sup>143</sup> Another CEC reported that they repaired school materials (chalkboards and chairs), in addition to repairing walls in a classroom.<sup>144</sup>

Where there is arguably the greatest disjuncture between the quantitative and qualitative findings concerns CEC provision of scholarships. As noted previously, just 7 of 63 head teachers surveyed at the midline reported that girls in their school receive scholarship support from the CEC. But financial support of individual students was a common theme in discussions with CEC members. Participants in FGDs noted that said that – while they cannot support all of the students who need it – they pay the school fees for many children.<sup>145</sup> Others indicated that as many as 70 students were currently attending school for free.<sup>146</sup>

One possible explanation for the discrepancy between head teachers' and CEC members' reporting of scholarship provision is in the terminology used and the precise nature of the financial support provided. When head teachers were asked about this topic, they were asked whether the CEC provides *scholarships* to students. When CEC members talked about their financial support for students, however, it was often framed in two ways that do not necessarily fit with the connotation of "scholarship": first, they mentioned paying school fees for students, occasionally in what seemed to a one-off or ad hoc manner.<sup>147</sup> For instance, a student whose family normally pays the cost of schooling but cannot afford school fees for a one month stretch may be supported by the CEC to remain enrolled, but this support may not be viewed as a scholarship, with its more formal and long-term connotation. Second, some schools allow students in need to enrol without paying the requisite fees. But this decision has consequences for the school's finances; to the extent that CEC members contribute money to ease the resulting shortfall – by paying teacher salaries or providing other financial support – they are indirectly paying students' school fees.

Broadly speaking, the analysis here paints a reasonably optimistic portrait of CECs and their financial engagement with the schools they oversee. Compared to baseline, more CECs pay a portion of teacher

---

<sup>141</sup> FGD – CEC Members

<sup>142</sup> FGD – CEC Members

<sup>143</sup> FGD – CEC Members

<sup>144</sup> FGD – CEC Members

<sup>145</sup> FGD – CEC Members

<sup>146</sup> FGD – CEC Members

<sup>147</sup> This distinction was also noted in our analysis of parental reports of "financial support to students," because this may include a wider range of CEC actions than "scholarship provision," including less formal, less routinized financial support than that implied by a "scholarship."

salaries, and the share of teacher salaries they pay has gone up.<sup>148</sup> Likewise, support for school repairs and the acquisition of supplies is fairly widespread. However, the qualitative interviews highlight two areas of ongoing concern.

First, CEC fundraising efforts are invariably intermittent. When asked whether they have conducted fundraising campaigns, CEC members discussed discrete campaigns undertaken for a specific purpose (such as filling a gap in teacher salaries, or repairing a classroom wall).<sup>149</sup> Similarly, when asked whether they plan to fundraise in the future, several CEC members indicated that they would do so if the need arose – saying for instance, “If we encounter more issues we will do fundraising.”<sup>150</sup> Only one CEC specifically stated that they plan to fundraise on a continuous (i.e. annual or semi-annual) basis, citing the school’s continual deficit.<sup>151</sup> Intermittent fundraising is problematic because it means that issues – such as teacher salary shortfalls – take longer to resolve and produce problems, such as discontent among teachers, until funds can be raised. Even more problematic is the assumption that schools will be able to raise funds in the short-term when a problem arises. The difficulty of fundraising varies over time, and future efforts may not be sufficient to cover unanticipated needs.

Second, CEC fundraising is geographically concentrated in a way that exacerbates local economic shocks. As several CEC members described, fundraising was necessary last year due to drought in their areas and the strain that this placed on household finances.<sup>152</sup> A greater number of students needed support for school fees and teachers also faced strain in their personal finances. However, the individuals from whom funds were sought – including CEC members themselves and community members with moderate financial means – were also facing hardship, which meant that fundraising was made more difficult. One CEC member said “members of the CEC were among the community members affected by the drought.”<sup>153</sup> CEC members and community members are subject to the same localized economic shocks as families with students in school, which means that fundraising is most difficult at the exact moment when it is most necessary.

Both the intermittent timing of fundraising and the localized nature of that fundraising makes it more difficult for schools to “smooth” their consumption – or expenses/outlays – and income over time. A more sustainable model of CEC financial engagement would use continual fundraising to create a small rainy day fund that could be drawn down in emergencies and replenished at other times.<sup>154</sup> Even more ambitiously, a rainy day fund that averaged risk over a large geographic area, especially across areas that are drought-prone in different periods, would be ideal. Such an arrangement may be difficult to organize, for both classic reasons of institutional design and due to the specific clan dynamics in Somalia. From an institutional design perspective, one community may be unwilling to subsidize another during their lean times if they are unsure whether the arrangement will last long enough for them to recoup their

---

<sup>148</sup> At the same time, this trend may suggest a weakening of official funding mechanisms through the government – i.e. that CECs have been forced to shoulder a larger share of the salary burden due to government retrenchment – which may weaken sustainability in the long run.

<sup>149</sup> FGD – CEC Members.

<sup>150</sup> FGD – CEC Members

<sup>151</sup> FGD – CEC Members

<sup>152</sup> FGD – CEC Members

<sup>153</sup> FGD – CEC Members

<sup>154</sup> This idea is analogous to what economists – in the context of fiscal policy – call countercyclical policy. In short, fundraising efforts and spending are conducted *against* local economic cycles: fundraising occurs during boom times, when fundraising is easiest, and the rainy day fund is drawn down during lean times, when fundraising would be most difficult.

costs. For instance, if subsidies flow primarily in one direction for the first several years of the arrangement and then the arrangement is dropped, the supporting community is unambiguously worse off. The design issue is exacerbated by clan dynamics: a rainy day fund that drew from more than a single clan may be fundamentally unworkable due to mistrust, and any disputes that arose might inflame communal tensions further. However, if institutional arrangements could be designed to mitigate this risk, a broad-based rainy day fund would help alleviate risk and make for more effective fundraising efforts by CECs.

Overall, CECs are engaged in more consistent financial support of their schools than they were one year ago during the baseline. CECs provide a greater share of teacher salaries presently, and a bare majority of CECs provide support for purchasing or repairing classroom materials and school infrastructure. CECs in intervention communities have also increased their monetary support of education in other ways from baseline to midline, including greater support for individual students and increased support for improving infrastructure. At the same time, there remain significant limitations to the ability of CECs to simultaneously pay for needed school improvements, top-up teacher salaries, and provide financial support to students in need. First, for many, the level of financial support they are able to provide is still limited – the majority of CECs do not provide teacher salary support, and many cannot fund infrastructure improvements. Providing support on all of these dimensions simultaneously is likely beyond the reach of all but the best-organized and best-resourced CECs. Second, fundraising efforts are intermittent and ad hoc, which contributes to a perpetual sense of crisis – funds must be raised quickly and at the worst possible time for local economic conditions. Continual and more extensive fundraising efforts would produce greater financial stability for schools, with all of the benefits that could bring, including improved teacher satisfaction, more consistent financial support for students in need, and an increased ability to plan infrastructure improvements over a longer period.

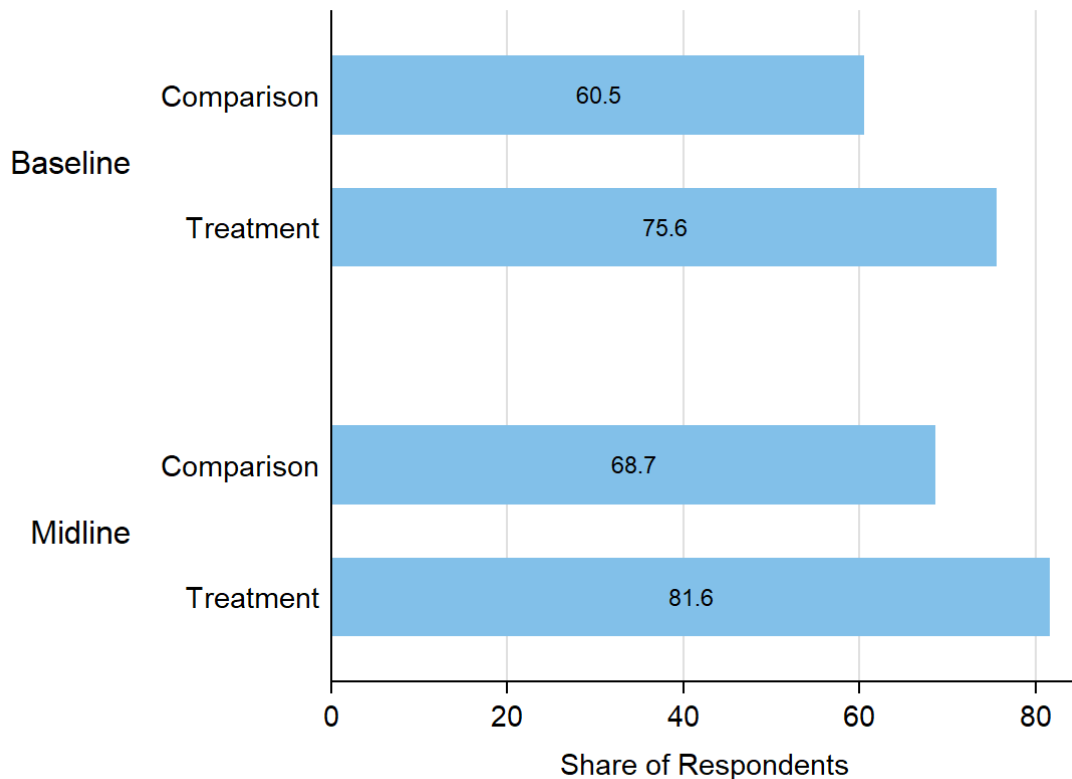
### **CEC Functionality and Activity Levels**

The second community-level sustainability indicator focuses on the extent to which CECs are perceived as functional by parents and community members. Given that CECs are responsible for a significant share of school oversight and management, parents should be broadly familiar with their activities, particularly if they are fulfilling their expected role.

Our main measure for this indicator is direct, asking parents whether their child's school has “a CEC that helps with school-related matters.” The indicator presents a fairly low bar, in that it does not specify an activity level that constitutes “active” or “functional”; rather, it asks whether a CEC exists and whether they help with school-related matters, large or small.

Overall, most parents report that their local school has a functional CEC, as shown in the figure below. Performance on this metric has improved since baseline in both intervention and comparison schools; in intervention schools, the share of parents reporting a functional CEC rose from 75.6 percent to 81.6 percent. While a similar improvement was observed among comparison schools, the purpose of this indicator is not necessarily to show impact vis-à-vis the comparison group, but to determine whether improvements made in intervention schools can be sustained following the end of the project. A 6-point increase in CEC functionality suggests progress toward sustainability.

Figure 16: Parents reporting that their school has a functional CEC



The improvements in parental perceptions are not limited to the mere existence and functionality of CECs. In fact, parents increasingly report that their CECs communicate with them regularly – either weekly or monthly – about their activities. Even among those parents who said their community has a CEC, the activity level of those CECs has increased: The share of parents in intervention areas who said their CEC either never communicated with them or did so only annually decreased from 28.2 and 8.9 percent at the baseline to 22.9 and 3.3 percent at the midline. Instead, the share of parents who said their CEC communicated with them monthly or weekly rose from 52.2 percent to 63.5 percent. Again, this understates the total improvement in intervention schools, it focuses exclusively on those parents who reported a functional CEC in the first place – an indicator that *also* rose among intervention schools from baseline to midline.

These findings fit closely with those derived from surveys of school head teachers and qualitative interviews with community members. In the case of head teachers, an increased share reported, at midline, that their schools had a School or CEC Management Plan. In intervention areas, the share of schools with such a plan rose from 43.8 percent at the baseline to 53.1 percent at the midline. Similarly, the share of head teachers who report that a CEC member has visited the school for monitoring purposes in the last year has increased from 59.4 to 87.5 percent from baseline to midline in intervention schools.

Meanwhile, qualitative interviews conducted with mothers at the midline seem to confirm the generally active role of CECs in their communities. When asked about CECs, mothers in every FGD acknowledged their actions; in most cases, their acknowledgement was seemingly unanimous, though mothers naturally



focused on different aspects or provided differing details about the actions their CECs took.<sup>155</sup> In those cases where CECs were not uniformly viewed as actively engaged in their communities, typically one mother dissented – citing either their own lack of knowledge about the CEC or the CEC’s lack of activity – while the remainder stated that the CEC was engaged with the local school.<sup>156</sup> As discussed in the context of CEC financial support for schools, CECs tend to be engaged most heavily in activities that do not require financial support of schools, such as awareness-raising, and school monitoring. This finding was corroborated by mothers in focus groups, who tended to emphasize the CEC’s role in community opinion formation, rather than their financial or material support of schools.

CECs are becoming more active in the management of their schools and more engaged in their communities. A small handful of intervention schools had less active CECs, where parents either reported very infrequent communication with the CEC, or where a significant minority or even a majority of parents stated that their community lacks a functional CEC altogether. Particular attention should be given to CECs in these communities, with more intensive monitoring of their activities, additional capacity-building and, perhaps, additional material support, if it is found that a lack of resources is reducing their community engagement.

### **Adherence to Implementation Standards for ALP, Numeracy and Remedial Classes**

This evaluation did not collect direct measures focused on implementation of ALP, numeracy and remedial class teaching standards. Our assessment of teaching quality tended to focus on teachers’ demeanour and interactions with students, the level of classroom participation, the use of corporal punishment, and so on, rather than adherence to specific standards, such as those set out in CARE’s teacher training programs. One area of implementation that the evaluation did capture was the use of formative assessment, which is a system of continual evaluation of student comprehension using a wide range of data points – e.g., interaction in the classroom, short assignments to gauge comprehension, or listening in on group work conversations. Formative assessments are widely considered a critical component of improved learning outcomes, as they help teachers identify students who are falling behind more quickly, and identify topic areas of particular concern.

CARE’s teacher training programs emphasize the use of formative assessments. We also study the use of formative assessments in our evaluation of teaching quality, in Section 5.3 of this report. Here we recap some of those results, because they are the best available indicator of adherence to implementation standards in SOMGEP-T schools.

The use of formative assessments has increased markedly since the baseline. Prior to the start of SOMGEP-T programming, approximately one quarter (26.2 percent) of teachers could show evidence of the use of formative assessments in their classroom. By the midline, this share of teachers had increased to 36.2 percent, with particularly steep improvements in intervention schools (increasing from 26.2 to 40.6 percent of teachers).

**Note from CARE:** It is important to note that the proportion of teachers using formative assessments described above refers to the teachers who claimed to use formative assessments, not those who could show evidence of use (records). The proportion of intervention teachers showing evidence of the use of formative assessments is much higher - 40% of the teachers observed in intervention schools could show records of formative assessments, against only 23% of those in comparison schools. The proportion of intervention teachers using formative assessments has also sharply increased in relation to the baseline (more details provided under the analysis of Intermediate Outcome 3).

---

<sup>155</sup> FGD – mothers

<sup>156</sup> FGD – mothers

Two caveats to this finding are important to note. First, the sample of teachers studied is very small – just 124 teachers at the baseline and 62 at the midline. More extensive data collection at the endline – especially a survey of teachers, as was conducted during the baseline – would increase our confidence in the identified trend. Second, teachers may overstate their use of formative assessments, either because they do not fully understand the term and its meaning, or because they are aware that the use of formative assessments is desirable, and they do not want to admit they do not use them. To partially address this issue, researchers also asked teachers whether they could provide evidence they used formative assessments, typically in the form of example short exercises or notes taken on student performance. Fewer teachers claimed this level of use; even here, however, the performance of intervention schools improved from baseline to midline, consistent with the trend in self-reported usage noted above.

### **Ministry of Education Engagement in Girls' Education Initiatives**

The responsibilities of the MoE include training CECs, recruiting teachers, monitoring classrooms and fundraising.<sup>157</sup> As noted at the baseline, Somalia's culture has generally preferred sending boys to schools while girls remain at home to help their family with chores. SOMGEP-T's intervention aims to support MoE staff (local education officers) in the following four areas: (1) Strengthening the capacity of Gender Departments to improve girls' education outcomes through trainings, development of action planning and provision of incentives to retain the gender focal points, particularly in rural areas, (2) Providing support to Regional Education Officers (REOs) and District Education Officers (DEOs) to mainstream improved teaching practices and address retention/transition issues (3) Working closely with MoEs' TVET/NFE Units to explore opportunities for vocational training, and (4) Advocating for employment of female pre-service graduates in target schools.

In the baseline evaluation, an analysis of information provided by regional education officers (REOs) was conducted in order to understand MoE performance. At the time of the baseline, performance varied—salary payments to teachers were often late or entirely lacking, in-kind support for schools was sporadic, and few MoE departments had gender units/departments. For the midline, we again conducted KIIs with REOs, and explored the performance of MoEs across the four dimensions mentioned above. Across a number of these dimensions, there appears to have been notable activity, but MoEs still struggle to consistently and uniformly provide support to schools, and activities appear to be sporadic and dependent on external assistance.

MoEs face a number of challenges that appear to be contributing to this dynamic, including financial constraints, lack of transportation, insecurity, lack of tools and training, and, in relation to implementation of this specific project, poor communication between project and MoE staff. In terms of financial support and in-kind contributions, MoEs are still severely constrained by lack of finances. In terms of financial constraints and transportation challenges, one REO explained, "Yes, we face some challenges, in particular financial challenges. This region is very large and sometimes you have to reach a place that can be 100km away. We only have one car available to us that belongs to the REO. Lack of finances also has an impact on petrol being available for the long journeys and will have an effect on whether we can reach far places and conduct supervisions."<sup>158</sup> Additionally, a number of REOs reported that MoE employees themselves often do not receive their salaries, and this problem of under-funding and unpaid salaries is consistent down to the level of teachers in local communities. As one REO explained, "The

---

<sup>157</sup> KII – REOs

<sup>158</sup> KII – REOs

economic situation of the country also affects the quality of education – some teachers, professors or even Ministry employees might not receive their salary for months or even years.”<sup>159</sup>

REOs across locations also mentioned challenges they face working with CARE International and other NGOs. First, lack of continuity in programming and delays in implementation appear to negatively impact MoEs’ plans. One REO explains, “I work at the Ministry of Education and in cooperation with SOMGEP. The project faces many challenges such as delays in implementation, and teachers do not receive a salary for even a full year. NGOs such as UNICEF, CARE, and Save the Children implement education-related projects but they also face many constraints and sometimes just stop implementation.”<sup>160</sup>

*Note from CARE: This remark reflects the interruptions in sector support projects (EU-funded Horumarinta Elmiga and Education Is Light; and GPE-funded ESPIGs) due to funding cycles. The projects have since then resumed operations.*

Another issue is poor communication between project staff and MoEs. This issue was raised by REOs from all three regions. One REO explains, “There is also a lack of communication. Some people do not know they will receive training and are not prepared for it.”<sup>161</sup> Another REO complains, “Organizations come for visits without informing us. Sometimes they inform us of a visit in a district very far from us, which we are unable to participate in under the time allotted. This undermines cooperation between us and CARE. I request that CARE informs us when they are working in the region so that we can witness the activities.”<sup>162</sup> Lastly, an REO interviewed explained that the project is not having its intended impact: “SOMGEP was helping 15 schools. The aim was to improve girls’ education and to support poor families. This has been going on for a long time and it did not cover their needs well. There is no change in the quality of education for the last two years because the plan did not take place as it was supposed to.”<sup>163</sup> This appears to be due to unmet expectations, as the REO explained that the region “did not receive support from the NGOs. The NGOs said they would pay for student fees but in the end they did not. I want to register this complaint to the NGOs.”<sup>164</sup> The REO also complains that CARE International was supposed to help the MoE organize more trainings, but says that “they did not work with us as we expected”<sup>165</sup> and as a result, the MoE was only able to hold one training in the past year.

*The proposed training, policy review, capacity building, grants for schools referred to by the REO is part of the plans by CARE’s Durable Solutions project and not SOMGEP T. CARE is currently implementing the Durable Solutions project funded by the EU Re-intake as part of the reintegration of refugees returning back to Somalia.*

Despite these constraints that hinder the ability of MoEs to reach schools consistently and uniformly, MoEs do appear to be engaging with schools to the extent possible. REOs reported that MoEs are providing a wide range of services. These services include conducting trainings, conducting supervision visits alongside CARE International’s staff, monitoring and evaluating education indicators, coordinating food provision programs for students, facilitating the work of NGOs, engaging with CECs and community members in awareness spreading activities, developing and in some cases implementing specialized

<sup>159</sup> KII – REO

<sup>160</sup> KII – REO

<sup>161</sup> KII – REO

<sup>162</sup> KII – REO

<sup>163</sup> KII – REO

<sup>164</sup> KII – REO

<sup>165</sup> KII – REO

plans for disabled students and pastoralists' children, directly supporting disadvantaged girls by covering school fees, and providing books and infrastructure support. This high level of in-kind programmatic support constitutes a major improvement over the baseline and will be discussed in more depth below.

#### *Establishment of a Dedicated Gender/Child Protection Unit Strengthening the Capacity of Gender Departments*

One of the core areas through which the project is providing support to MoEs is in strengthening the capacity of Gender Departments to improve girls' education outcomes through trainings, development of action planning, and provision of incentives to retain the gender focal points, particularly in rural areas. In only two out of the six REO interviews, respondents explicitly confirmed that there is a gender unit or department, but respondents from other interviews mentioned conducting activities that are consistent with the activities of a dedicated gender unit.

Among the activities mentioned by respondents were the construction of separate toilets for girls, provision of specific trainings on gender equality, distribution of sanitary kits, and promotion of female teachers to encourage girls' enrolment and retention. However, progress in this area appears to still be highly dependent on outside assistance, as one REO explained: "The MoE has a gender office and every region has a female member. However, the MoE does not have the financial means to pay these people. The Ministry only receives help from CARE and thanks to that we have been able to improve."<sup>166</sup>

#### *REOs and DEOs Mainstreaming Improved Teaching Practices and Addressing Retention and Transition Issues*

Another core area of project support is in working with REOs and DEOs to mainstream improved teaching practices and address retention/transition issues. One way in which REOs and DEOs are addressing girls' retention and transition issues is through the promotion of female teachers, as mentioned above. One REO directly links dropping out to lack of female teachers: "Female students also feel shy to ask the male teacher to repeat the lesson if they do not understand it. This can cause the girls to drop out of school because they feel left out."<sup>167</sup> In recognition of this problem, REOs are focusing on recruiting female teachers: "Yes, one of the things is to try to increase the number of female teachers in order to attract more female students."<sup>168</sup> Additionally, MoEs have made supervision visits, at times alongside CARE International staff, in order to track education metrics: "The project has been very useful because they created awareness and started supervising and managing the school better. We also monitor closely the educational situation of girls..."<sup>169</sup>

The qualitative data also provides evidence of progress made in improving teaching quality. A number of REOs mentioned that special efforts have been made to recruit high quality teachers and to train teachers on how to teach difficult subjects and work with disadvantaged children. One REO describes working with NGOs to identify areas in which teachers need improvement: "Yes, we found out that the subjects of math and English were difficult for the students and we shared our feedback with the NGOs. The teachers then received more training in these subjects in order to increase their knowledge."<sup>170</sup> Another explains, "We evaluate the number girls that drop out of school, because special teachers have been trained and CARE pays their school fees. They evaluate the registration of the school and support with the syllabus even if

---

<sup>166</sup> KII – REO

<sup>167</sup> KII – REO

<sup>168</sup> KII – REO

<sup>169</sup> KII – REO

<sup>170</sup> KII – REO

they cannot afford to.”<sup>171</sup> Teachers have also been trained in some locations on proper treatment of children with disabilities, and in one location, there is a mobile teacher who teaches pastoralist children.

However, there has not been uniform progress made in improving teaching quality and addressing retention and transition issues across locations. One quote in particular captures the reason why there is high variation in the quality of education across locations: “The quality of education always depends on the school management or the school principal. Education is like a chain, commands from the district to the regional level are interrelated and affect it. The economic situation of the country also affects the quality of education – some teachers, professors or even Ministry employees might not receive their salary for months or even years. We face many challenges. For example, school curriculums are incomplete, or there are security issues constraining the provision of education in many regions.”<sup>172</sup>

Other REOs explained the specific challenges faced in their areas. As one explains, “Another challenge is that districts are made out of lots of small villages that do not have the finances to recruit a good qualified teacher. What they then do is appoint someone that is from their village that does not have any qualifications or a lot of knowledge.”<sup>173</sup> Whereas in some areas, specialized plans have been developed for children with disabilities, in other areas, MoEs are unable to meet their needs: “We do not have the ability to fulfil the needs of disabled children living in the countryside. We built a school for children with disabilities. There are between 60-70 children, including deaf children, with a disability at the school and NGOs do not contribute to the school.”<sup>174</sup> The same REO described how plans to address the unique challenges of pastoralist children have not been fully implemented: “We formed a scholarship for pastoralist children to attend school but we still have not brought teachers out to the countryside.”<sup>175</sup>

#### *TVET/NFE Units Exploring Options for Vocational Training*

The project is also working closely with MoEs’ TVET/NFE Units to explore opportunities for vocational training. There is evidence from the qualitative interviews that MoEs are formulating, and in some cases implementing, plans to provide vocational training to drop-outs and adults. One REO explained that, “The ones that dropped out and we managed to track down, we provide them with vocational training skills.”<sup>176</sup> Another explained, “Right now our plan is to teach mothers technical vocational skills they can work with like henna, tailoring, cooking and business. Those skills enable the families to stay in the villages and not move anywhere else because there are a lot of people whose livestock were killed by drought.”<sup>177</sup> An REO suggested that the MoE is paying special attention to the challenges facing adult women: “One of the main challenges facing girls is adult education for women. Since the civil war affected the country, women had an incomplete education. Actually, for both girls and boys.”<sup>178</sup> In another location, an REO reports that the ministry built a school for adults in the community and provides vocational trainings for the unemployed and those who have aged out of the school system.

#### *Advocating for Employment of Female Pre-service Graduates*

Lastly, the project is working with MoEs to advocate for the employment of female pre-service graduates in target schools. There is some evidence from the qualitative interviews that MoEs are establishing programs that encourage employment of female graduates, but the evidence is sparse. For example, at least one region has established scholarships for girls to go to university so that “they will learn something and return

---

<sup>171</sup> KII – REO

<sup>172</sup> KII – REO

<sup>173</sup> KII – REO

<sup>174</sup> KII – REO

<sup>175</sup> KII – REO

<sup>176</sup> KII – REO

<sup>177</sup> KII – REO

<sup>178</sup> KII – REO

their region.”<sup>179</sup> In another region, the REO explains that “the students that have finished secondary education get to volunteer somewhere for one year.”<sup>180</sup> However, further detail on these points was not provided, and the sparse evidence suggests that this is not currently a priority issue for MoEs, likely because they are already struggling with a lack of finances in other priority areas.

#### *Triangulation with Quantitative Data*

Finally, the analysis of MoE engagement triangulates system level data, specifically the data gathered from schools and primary caregivers of cohort girls. MoEs are responsible for training CEC members, and, as such, the quality of CEC engagement with schools and communities serves as an indicator of the quality of MoE engagement. The analysis below uses the 63 schools that were part of both the baseline and midline studies and thus constitute the comparable longitudinal sample of schools. At the school level, the prevalence of CEC management plans and the level of CEC activity among sampled schools are indicators, and at the community level, the caregivers’ perception of CEC support to schools and caregivers’ family involvement in CECs serve as indicators of MoE engagement.

At the baseline, a relatively small percentage (28.6 percent) of schools reported having a CEC management plan.<sup>181</sup> That percentage has increased substantially to 46.0 percent at the midline. Meanwhile, the percentage of comparable schools (head teachers) that reported receiving a monitoring visit from a CEC member increased from 41.3 percent at the baseline to 77.8 percent at the midline. These findings suggest that CEC engagement has increased substantially since the baseline.

Despite the fact that these school-level measures are moving in the direction of increased CEC engagement, an analysis of change over time in the intervention group versus change in the comparison group suggests that the intervention has probably not been responsible for the observed increase in engagement. The table below summarises CEC engagement indicators by baseline versus midline and intervention versus control, showing the net difference in differences in the final column of the table. The difference in differences are negative for both indicators because the net increase in the proportion of comparison schools with CEC engagement from baseline to midline was larger than the net increase in the proportion of intervention schools. Neither of these difference in differences are statistically significant, so all that can be concluded from this analysis is that there is no detectable positive effect of the intervention in terms of increasing CEC engagement as measured through school-level indicators.

*Table 47: School-level indicators of CEC engagement*

Indicator	Baseline Intervention Schools	Midline Intervention Schools	Baseline Diff	Baseline Comparison Schools	Midline Comparison Schools	Comparison Diff	Diff in diff
School has CEC management plan	43.7%	53.1%	9.4%	12.9%	38.7%	25.8%	-16.4%
School received monitoring visit from a CEC member in the past year	59.4%	87.5%	28.1%	22.6%	67.7%	45.1%	-17%

<sup>179</sup> KII – REO

<sup>180</sup> KII – REO

<sup>181</sup> Note that the figure in the baseline report is slightly different, but this figure has been updated such that it reflects only schools that were included in both the baseline and midline samples and are thus comparable.



At the community level, the majority (68.6 percent) of primary caregivers at the baseline reported that their local school has a CEC which provides support for school-related matters, and this increased to 75.7 percent at the midline. In comparison, community participation in CECs at the baseline was relatively weak, with only 18.6 percent of sampled caregivers reporting that they were involved in a CEC or have a family member who participated in CEC activities, and that proportion slightly decreased at the midline (17 percent).

In keeping with the findings related to school-level indicators above, there is no evidence that the intervention has had a positive effect on CEC engagement as observed at the community level. The table below summarises community-level indicators by baseline versus midline and intervention versus comparison, with the difference in differences presented in the final column. Since the baseline, there was an increase in the proportion of caretakers reporting that their child’s school had a CEC, but the increase in the comparison group was larger than the increase in the intervention group, giving a negative difference in differences. The reasons for these substantial improvements over time in comparison schools are not clear from the available evidence, but it is worth noting that the baseline proportions for comparison schools were particularly low at baseline (much lower than intervention schools), which permitted far greater room for improvement. When it comes to CEC involvement, the proportion of involved family members increased slightly in the intervention group since the baseline and decreased slightly in the comparison group, resulting in a positive, but small difference in differences. Ultimately neither of the two community-level indicators are statistically significant in terms of their difference in differences. Thus, at the community-level as well, there is no detectable positive effect of the intervention in terms of CEC engagement.

*Table 48: Community-level indicators of CEC engagement*

Indicator	Baseline Intervent. Caretakers	Midline Intervent. Caretakers	Intervent. Diff.	Baseline Comparison Caretakers	Midline Comparison Caretakers	Comparison Diff	Diff in Diff
School has a CEC that helps with school-related matters	75.6	81.6	6	60.6	68.7	8.1	-2.1
Caregiver or member of family is involved in CEC	12.2	13.3	1.1	13.4	12.3	-1.1	2.2

By both qualitative and quantitative measures, the intervention does not appear to have increased MoE engagement with schools and CECs. Qualitative evidence confirms that the MoE is making a variety of significant in-kind contributions at the school-level, but financial contributions are likely to be the most impactful, and these are also comparatively rare as resources are exceedingly limited at every level from the Ministry to the community. At the school-level, number of active CECs has increased substantially at the midline, including the number of CECs who are engaged in monitoring schools; however, these increases have been just as common among comparison schools as among intervention schools, suggesting that CEC activism is not being driven by the intervention itself.



## **Inclusion of ALP in National Non-Formal Education Frameworks**

Qualitative evidence from MoE officials and REOs suggests that formal recognition of ALP is minimal, but that awareness is building now that these interventions have begun. Only one REO explicitly mentioned the ALP program by name, explaining that ALP programs are now functioning as a result of the project: “Previously the number of girls attending school was very small but thanks to the encouragement and payment of fees by CARE International, the number has increased significantly. Girls who dropped out of school are now attending the ALP program. Old people are receiving informal education. There is now a loan association group.” However, the responses of a number of other REOs suggest that alternative learning programs are being implemented in a variety of forms. One REO described how the MoE in the area is specifically targeting pastoralists’ children: “They made a mobile teacher that travels with the nomad family and also travels between the families and it was very beneficial.”

## **6. Key Intermediate Outcome Findings**

### **6.1. Attendance**

Improving girls’ attendance is a fundamental intermediate outcome of the project. According to the project’s Theory of Change, which was empirically tested and validated at the baseline, there is a positive relationship between attendance and student learning outcomes. Using a difference-in-difference approach, the findings in this section do not determine that the attendance of the in-school girls in the intervention group has significantly improved since the baseline compared to their counterparts in the comparison group.

The findings on attendance were derived from multiple measures taken in the headcount survey, the school survey, and the household survey. The findings from the attendance measure in each survey are then triangulated and compared.

#### **Attendance from Headcount Survey**

In the headcount survey, enumerators visited schools and collected data through two methods: 1) recording students’ attendance from the school’s attendance register for the day before the visit and the day of the visit and 2) performing a direct headcount of students on the day of the visit. Enumerators started collecting data one hour after the beginning of the classes until one hour before the lunch break. This provided the teachers an adequate amount of time to record attendance and collect data on students who may only attend school for half of the day. Excepting pre-kindergarten and kindergarten classes, all other classes in the school were included in the headcount. The list of the headcount survey questions is listed below for reference:

*Headcount Survey Questions*

- B7. Enter the number of GIRLS enrolled in this class
- B9. Teacher count on record: Number of girls marked in class YESTERDAY
- B10. Teacher count on record: Number of girls marked in class TODAY.
- B11. Girls HEAD COUNT in class (done by Enumerator): Enter the total number of GIRLS present in the class by counting
- B12. Enter the number of BOYS enrolled in this class
- B14. Teacher count on record: Number of boys marked in class YESTERDAY
- B15. Teacher count on record: Number of boys marked in class TODAY
- B16. Boys HEAD COUNT in class (done by Enumerator): Enter the total number of BOYS present in the class by counting

As shown in the table below, both girls and boys had slightly lower attendance rates in the midline than in the baseline along all measures of attendance. In addition, the headcount rates are lower than the rates collected from attendance registers both the day before and the day of the classroom visit by our enumerators. On average, girls had a headcount of 81.5 percent at midline, and register attendance rates of 81.7 percent on the day of the visit and 83.2 percent the day before the visit. The decline of attendance rates as the measure of attendance rate becomes more immediate in both the baseline and midline data suggests that attendance rate collection continues to be irregular and may have been completed selectively to improve perceptions of the school.

Girls' attendance rates were generally lower than the boys' attendance rates: among all enrolled boys, 80.9 percent were present during the headcount, 83.5 percent were recorded as present in the register on the day of the visit and 85.5 percent in the register on the day before the visit. The main finding from this analysis is that there is not a statistically significant attendance gap between girls and boys at the midline. There was no gap at baseline either, suggesting that, in the aggregate, girls are no more likely to miss school than their male peers.

*Table 49: Attendance Yesterday, Today, and of Headcount for Girls and Boys in Midline and Baseline – Headcount Survey*

Attendance (Baseline)	Attendance		Attendance (Midline)	Attendance	
	Girls	Boys		Girls	Boys
Register Yesterday (%)	87.5 (n=280)	89.7 (n=276)	Register Yesterday (%)	83.2 (n=281)	85.5 (n=278)
Register Today (%)	84.4 (n=323)	84.2 (n=318)	Register Today (%)	81.7 (n=299)	83.5 (n=279)
Headcount today (%)	82.8 (n=392)	81.8 (n=378)	Headcount today (%)	81.5 (n=420)	80.9 (n=412)

The majority of qualitative interview participants report that girls' education indicators have improved and attribute positive changes to the project and community efforts. An REO explains this improvement: "In our region four students were girls among the total six students...I believe the improvement of the quality of the girls doubled compared to the boys as I have told you and this was a result of the hard work from

CARE International, teachers, school administration and the parents of the students.”<sup>182</sup> Another REO explains, “The community is providing more support as they are now interested about education. The committee members have received training and guidance and have brought their ideas to the school. Girls are now aware of the importance of continuing education instead of pursuing an early marriage. Therefore, the number of girls enrolled at school has increased considerably and they show motivation to continue their education.”<sup>183</sup> In a number of cases, qualitative interview participants actually noted that more attention should be paid to boys’ education, as girls’ enrolment and performance has in some case surpassed that of boys.

Turning to the question of how attendance has changed since the baseline, the tables below summarise changes over time in boys’ attendance rates and in girls’ attendance rates for each of the three measures of attendance and show attendance by baseline versus midline and intervention versus comparison. The final column in each table shows the difference in differences when comparing change over time in the attendance rates of intervention schools versus change over time in comparison schools.

*Table 50: Boys’ Average Attendance Rate – Headcount Survey*

Indicator	Baseline Intervention Schools	Midline Intervention Schools	Intervention Diff	Baseline Comparison Schools	Midline Comparison Schools	Comparison Diff	Diff in diff
<b>Register Yesterday</b>	89.61	84.79	-4.82	89.95	86.83	-3.12	-1.70
<b>Register Today</b>	84.18	82.51	-1.67	84.41	84.96	0.55	-2.22
<b>Headcount Today</b>	80.91	82.91	2	82.79	79.09	-3.7	5.7

*Table 51: Girls’ Average Attendance Rate – Headcount Survey*

Indicator	Baseline Intervention Schools	Midline Intervention Schools	Intervention Diff	Baseline Comparison Schools	Midline Comparison Schools	Comparison Diff	Diff in diff
<b>Register Yesterday</b>	88.78	82.99	-5.79	86.16	83.78	-2.38	-3.41
<b>Register Today</b>	84.72	80.46	-4.26	84.16	83.44	-0.72	-3.54
<b>Headcount Today</b>	81.89	82.04	0.15	84.01	80.95	-3.06	3.21

The main finding that emerges is that only the headcount rates of intervention schools show signs of improvement since the baseline.

<sup>182</sup> KII – REO

<sup>183</sup> KII – REO

The attendance rates of both boys and girls gathered from the attendance records the day before or the day of the visit changed only marginally and the changes were largely negative. Moreover, the declines in recorded attendance rates were often greater in the intervention schools than in the comparison schools, and therefore, these measures' net difference in differences is negative. For the headcount attendance, the students in the intervention group had greater attendance compared to their counterparts in the comparison group. Positive net differences were observed among the intervention group from baseline to midline, and negative net differences were observed among the comparison group over time resulting in positive difference in differences. Ultimately, none of the difference in differences are statistically significant, suggesting that the project has not had a measurable impact on attendance during the four months of exposure to the intervention. The percentage difference in the average attendance rate at the midline compared to baseline across the intervention and comparison group is indicated in the table below.: Percentage Difference in the Average Attendance Rate at Midline vs. Baseline

*Table 52: Percentage Difference in the Average Attendance Rate at Midline vs. Baseline*

		Intervention	Comparison
Yesterday	Boys	5.3 ↓	3.4 ↓
	Girls	6.5 ↓	2.7 ↓
Today	Boys	1.9 ↓	0.6 ↑
	Girls	5.0 ↓	0.8 ↓
Headcount today	Boys	2.4 ↑	4.4 ↓
	Girls	0.2 ↑	3.6 ↓

For other important subgroupings such as grade-level, the only statistically significant difference in differences observed was among boys in grade 3 (3) whose average headcount rate has increased from 75.7 at the baseline to 82.4 at the midline in the intervention group compared to their counterparts in the comparison group whose average headcount rate show a decrease from 83.4 to 71.5. Nevertheless, the intervention group of boys in grade 7 (8) have had lower average headcount rate as well as today's attendance rate compared to the comparison group. More specifically, the intervention group's average headcount rate has dropped from 91.0 to 81.6 while the comparison group's rate has improved from 82.0 to 90.2 for the same period. Data shows a similar decline rate in the intervention group's average today's attendance rate with 93.8 at the baseline to 82.5 at the midline whereas the comparison group's today's rate boosted from 80.4 to 90.1.

To briefly summarise the key findings above, only the boys' headcount rate in grade 3 (3) has measurably increased since the baseline as a result of the intervention. On average, headcount attendance rates have slightly increased while the yesterday's and today's attendance rates have decreased very slightly over time for both boys and girls.

### **Attendance from School Survey**

As in the baseline, in the school survey, survey team leaders were asked to obtain enrolment and attendance records from the head teacher or principal of each school. Survey team leaders were then asked to record the enrolment status and attendance records of each girl who was in school in the baseline.

### School Survey Questions

J10. How many days has GIRL attended during this school year (so far)?

J11. Record: number of possible attendance days so far this year

The mean attendance rate of cohort girls, already relatively high, has not increased significantly from baseline to midline, however substantial gains have been made in improving the attendance recordkeeping between the rounds. The mean attendance rate of 541 girls for whom schools had school attendance records in the midline was 92.8 percent. The 541 girls who have attendance records compose 70.3 percent of the total cohort girls sampled in this round. The midline mean attendance rate of cohort girls is not a significant improvement on the attendance rate of girls in the baseline, 93.0 percent, which was collected from the records of 392 cohort girls who composed less than half, 48.9 percent, of all cohort girls.

Table 53: Girls' Attendance in Baseline and Midline – School Survey

	Baseline		Midline	
	Intervention	Comparison	Intervention	Comparison
Mean attendance rate (%)	91.9	94.5	91.7	94.3
Girls with attendance records (%)	52.1	45.7	76.9	62.7
Total number of cohort girls	428	370	428	370

The improvement in attendance recordkeeping of girls was a significant improvement from baseline to midline.<sup>184</sup> In addition, while 26 of 64 schools in the baseline had no records for any of the cohort girls sampled, in the midline, only 7 of 64 schools continued to have the same problem.

In the qualitative interviews, teachers across regions reported that attendance is formally tracked using registers/attendance sheets. A teacher explains that teachers are now able to follow students' movements as a result of increased efforts to track attendance: "The school has a register, so every teacher uses his register and from it he will know who is absent and who is present. So, the change is that the teacher can follow the movements of the students."<sup>185</sup> In some cases, teachers follow up with parents when students are absent, and in others, the principal or CEC members keeps track of attendance. A teacher explains: "Students' attendance is being tracked daily, except on Fridays. You can track every student who is present or absent, and then the information gets shared with their parents to ask about the reasons for their absence."<sup>186</sup> In other cases, students are given menial tasks, such as sweeping the floor, as punishment for absence, or teachers "inform the students they will lose marks if they miss the class."<sup>187</sup> However, there is evidence that teachers handle these issues on a case-by-case basis and that special allowances are made for students who live far or whose parents are pastoralists. As one teacher explains, "Students are not the same and come from different distances. Some students need to travel up to three hours to get to school while others only have to travel less than an hour to get to school. We understand and accommodate for the individual circumstances of each student."<sup>188</sup>

<sup>184</sup> The finding is significant with a P-value of 0.003 using cluster robust logistic regression.

<sup>185</sup> FGDs – Teachers

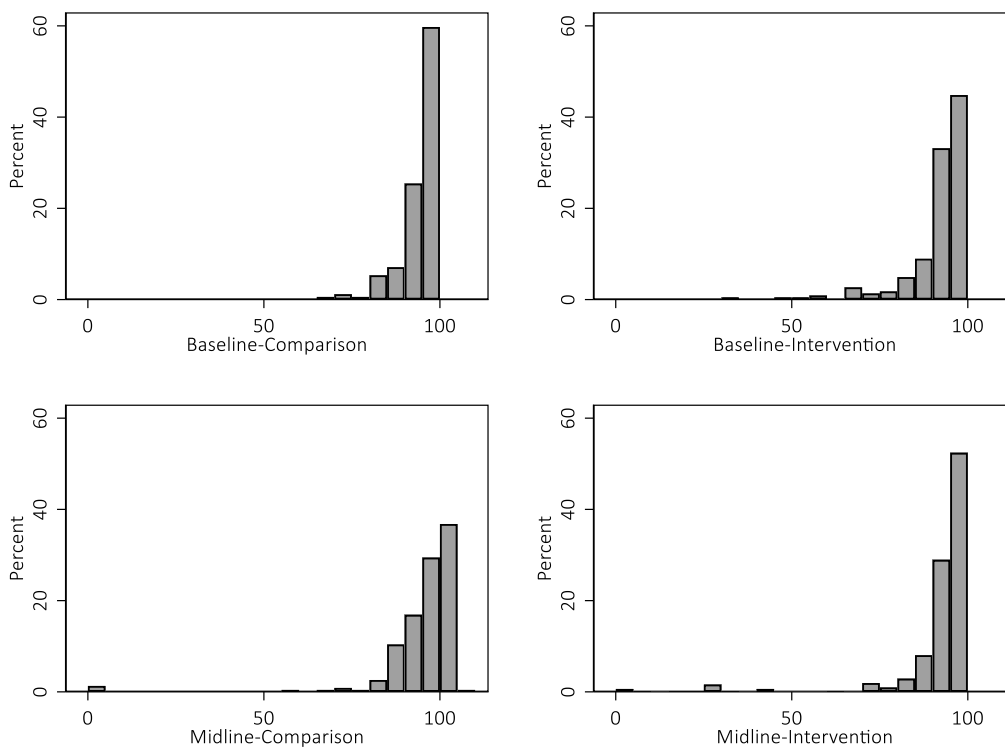
<sup>186</sup> FGDs – Teachers

<sup>187</sup> FGDs – Teachers

<sup>188</sup> FGDs – Teachers

The histograms below illustrate the percent of the sample with a given rate of attendance in the baseline as well as in the midline. The majority of cohort girls in both the baseline and the midline have high levels of attendance as seen in the left-tailed histograms below.

*Figure 17: Girls' Attendance in Baseline and Midline – School Survey*



Nevertheless, substantial percentages of girls have missed sizable portions of their school year. The table below presents the proportion of girls by decreasing rates of attendance. While the girls in the midline are somewhat less likely to have a limited level of attendance than baseline girls, the gains in the midline are uneven. Cohort girls in the midline were less likely than in the baseline to have an attendance rate less than 90 percent (17.2 percent vs. 18.9 percent) or 85 percent (8.4 percent vs 10.7 percent). However, the difference in the proportion of girls who attended less than 80 percent of the time between the baseline and the midline is marginal (5.7 percent vs 5.6 percent). Midline cohort girls are also more likely to have girls with very low levels of attendance: 3.2 percent of cohort girls in the midline have attendance rates less than 60 percent compared with 1.3 percent of cohort girls in the baseline.

*Table 54: Limited Girls' Attendance in Baseline and Midline – School Survey*

Level of Attendance	Baseline		Midline	
	Intervention	Comparison	Intervention	Comparison
<90% attendance	22.0%	14.8%	17.7%	16.5%
<85% attendance	13.0%	7.7%	10.1%	6.1%
<80% attendance	8.1%	2.4%	7.3%	3.5%
Total number of cohort girls	223	169	328	231

## Attendance from Household Survey

### *Household Survey Questions*

PCG\_5enr. Since the start of the most recent school year, has GIRL attended her (main) school on most days that the school was open?

PCG\_6enr. Has she attended more than half the time, about half the time, or less than half the time?

PCG\_6enr\_na. How many days of schooling did she miss last month?

In the interview with primary caregivers within the household survey, primary caregivers are asked a number of questions about their cohort girls' school attendance.

Of the 764 primary caregivers of in-school girls who were asked these questions, 89.1 percent said that their cohort girl attended most days since the start of the school year. Among those who said that their cohort girl did not attend most days, 54.9 percent said that their girl attended more than half the time, 22.0 percent said she attended about half the time, and 12.2 percent said she attended less than half the time. Significant differences in the responses between intervention and comparison groups and between the baseline and midline were not observed.

Qualitative interview respondents shared a number of reasons why girls might be missing school either some of or most of the time. One of the most commonly mentioned reasons was sickness, although this is likely because the only girls who were directly questioned about their attendance in the qualitative portion of the midline research were girls with disabilities. However, girls with disabilities did not *only* talk about personal illness or difficulties associated with their disability preventing them from going to school – they also in some cases shared that they have missed school as a result of an illness or death in the family. As one girl explains, “I also miss school when my mother is sick. There was another time that I missed school because one of my brothers died.”<sup>189</sup> A few mothers also mentioned sickness as a cause of absence, with one stating that “falling ill – either the girl or the mother or either one of them”<sup>190</sup> is one of the main reasons a girl might not attend school and another stating that sickness is one of the “only things that can stop girls from attending school.”<sup>191</sup> It should be noted that “sickness” is oftentimes the term used to refer to a girl who is on her period. One mother explained that girls “do not attend school when they are scared or when they have their period,” which may in some cases be due to a lack of private toilets, as this was mentioned by an REO as a reason why girls might miss school.

A number of respondents also mentioned parents keeping their girls home to help out with work or household chores. One girl explained that she stays home when she has to help her mother “look after her shop,” another that she sometimes misses school to help her mother “with herding livestock,” and yet another that she misses school whenever her mother asks her “to do something.” This likely refers to household chores, as respondents also explicitly mentioned household chores as a major issue with regular attendance. One mother explained that “if a mother does not have enough children to help at home activities that mother might instruct her daughter to stay at home and help the family.”<sup>192</sup> In a KII, an REO explained, “Girls are often late to school due to doing household chores.” Mothers in one area listed

---

<sup>189</sup> IDI – GwD

<sup>190</sup> FGD – Mothers

<sup>191</sup> FGD – Mothers

<sup>192</sup> FGD – Mothers



household chores as a major reason why girls might even be forced to drop out of school: “Marriage and having to do house chores are the main reasons why girls stop attending school.”

Other reasons qualitative respondents mentioned that girls might not attend school were insecurity and drought. Again, it should be noted that insecurity was mentioned as a reason for drops in attendance only by respondents from an area that recently experienced inter-clan conflict; as one mother explains, “The school remained closed for six months because we were facing insecurity.”<sup>193</sup> Drought, however, was mentioned by respondents across areas – one teacher shared that “student absence increased during the drought, but it was normal under other circumstances,”<sup>194</sup> and a mother from a different area provided a potential explanation for this observed decrease in attendance in sharing that “the quality of education decreased during the drought,” and that “parents sent their children to collect jerry cans instead of attending school.”<sup>195</sup>

In some cases, distance to school may also be a major barrier to attendance, particularly during times of hardship or in certain seasons. One mother explained, “The secondary school is located far away and because of that students will arrive home late, which is a problem.” A teacher from a different area also cited distance as an issue for some students: “Students are not the same and come from different distances. Some students need to travel up to three hours to get to school while others only have to travel less than an hour to get to school. We understand and accommodate for the individual circumstances of each student.” A teacher from this same area explained that attendance increases during the rainy season, as “most of the community members who live here are pastoralists.”<sup>196</sup>

Based on the number of days of schooling the primary caregiver said that the cohort girl missed in the last month, a rough estimate of the cohort girl’s attendance rate was also calculated.<sup>197</sup> The estimated average attendance rate of all in-school cohort girls in the midline is 90.9 percent. This is only a marginal improvement on the attendance rates estimated in the baseline, 90.3 percent. Although the increases in attendance rates from baseline to midline were not statistically significant, the rate of attendance rose among both intervention (90.7 percent to 91.2 percent) and comparison groups (89.8 percent to 90.7 percent).

As with attendance rates calculated through the headcount survey and the school survey, attendance rates estimated from the household survey are left-tailed as shown in the figures below. The distribution of attendance rates are similar from baseline to midline, but for both comparison and intervention areas the attendance rates have shifted moderately to the right.

---

<sup>193</sup> FGD – Mothers

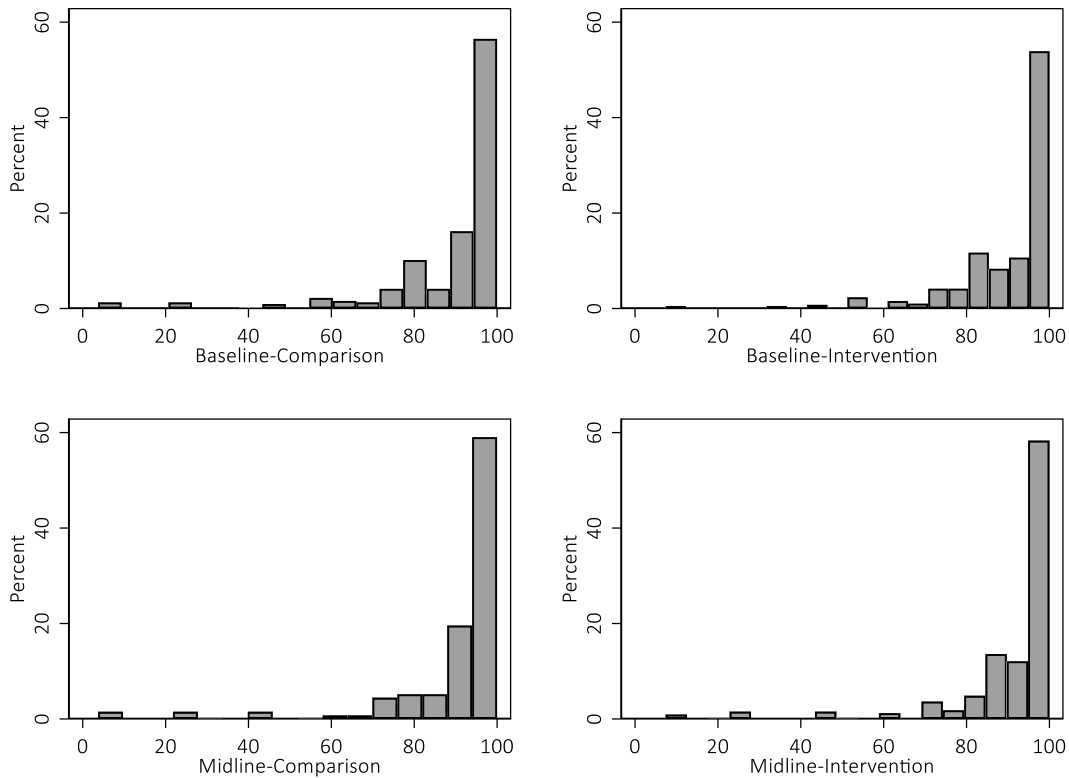
<sup>194</sup> FGD – Teachers

<sup>195</sup> FGD – Mothers

<sup>196</sup> FGD – Teachers

<sup>197</sup> The number of days missed was subtracted from 27 and then divided by 27 to arrive at the estimated attendance rate. Somali girls go to school except for every Friday of the month. Given that there are 4 Friday’s each month and the maximum number of days is 31, 27 was assumed to be the maximum number of days a girl could attend school. This estimate will slightly underestimate the attendance rates if the primary caregiver assumed a total number of days less than 31 when she was thinking back to last month.

Figure 18: Girls' Attendance in Baseline and Midline – Household Survey



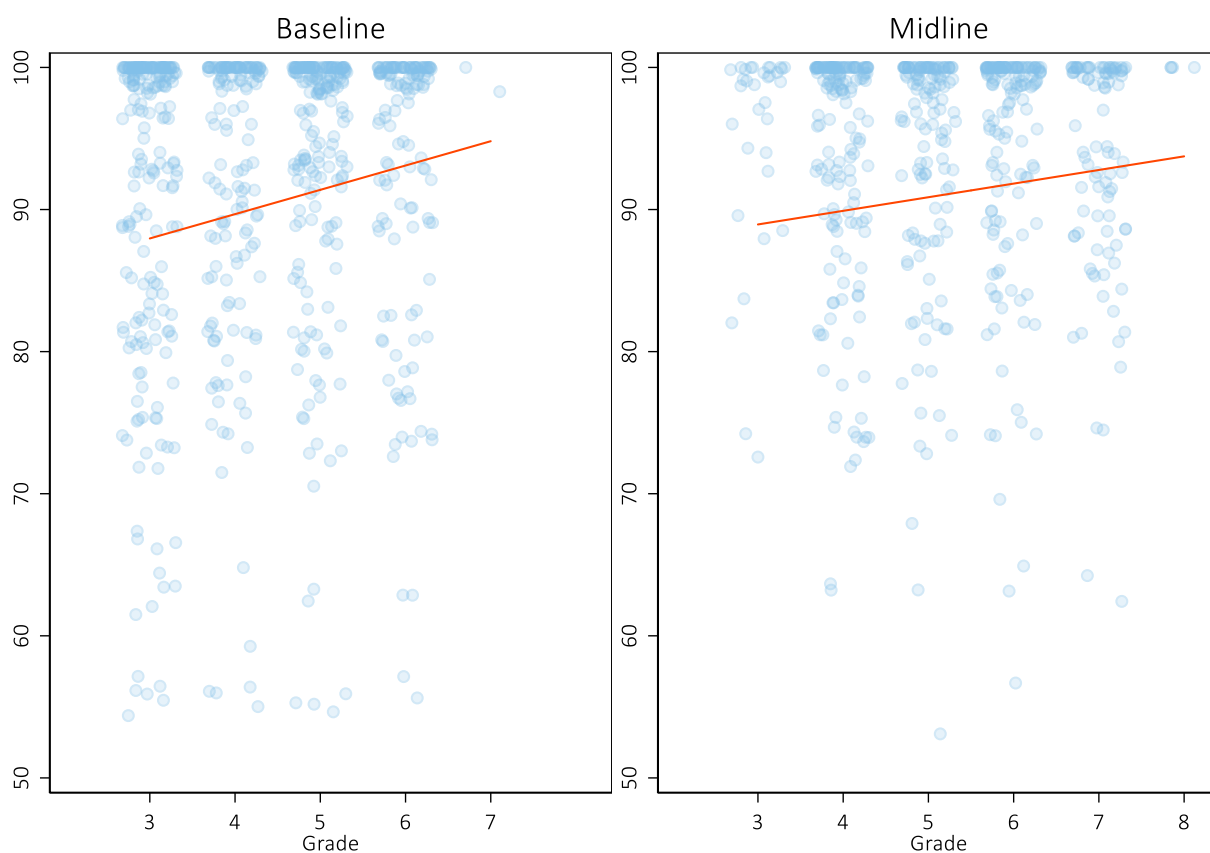
Correspondingly, while not significant, the proportion of girls who have lower attendance rates have decreased, as shown in the table below. Girls who were estimated to have attended less than 90 percent of school days composed 30.2 percent of the sample of in-school cohort girls, 16.1 percent attended less than 85 percent of the time, and 11.6 percent attended less than 80 percent of school days last month. The decrease in proportion of girls with limited attendance from baseline to midline was observed in the comparison and intervention groups, and the difference in difference was not significant.

Table 55: Limited Girls' Attendance in Baseline and Midline – Household Survey

Level of Attendance	Baseline		Midline	
	Intervention	Comparison	Intervention	Comparison
<90% attendance	35.4%	33.3%	29.6%	31.0%
<85% attendance	23.7%	23.2%	16.0%	16.2%
<80% attendance	15.4%	16.8%	11.2%	12.2%
Total number of cohort girls	384	315	331	271

In the baseline there were significant differences observed by grade, with increasing rates of attendance as grade levels increased. While attendance rates rise with each additional grade in the midline data, as shown in the figure below, this effect is not significant.

Figure 19: Girls' Attendance by Grade – Household Survey



### Triangulation of Attendance Rate Findings

The three baseline attendance rates obtained in the above analysis reflect findings for three timeframes. The attendance rate gathered from the headcount conducted by the enumerator in the headcount survey reflects the most accurate attendance rate since it was collected by a third party in the classroom, but it is the most limited in terms of generalisability, because it only offers a snapshot of attendance on the day the survey team visited a school. The attendance rate from the household is gathered over the past month, but it is reported in terms of the number of days the girls were absent in the past month and relies on the memory of the primary caregiver being interviewed. The attendance rate of the school survey covers the broadest period of time, the academic year thus far (approximately two months at the beginning of fieldwork), but relies on attendance records that are frequently incomplete even as there has been a significant improvement in attendance recordkeeping.

In the baseline, our analysis found that the less immediate a measure of girl school attendance is, the higher the average attendance rate tends to be, and in the midline this trend continues as shown in the table below which presents the attendance rates gathered from each of the survey tools across rounds as well as intervention and comparison groups within each round. The most proximate measure of attendance, based on the headcount of student in class on the day of the visit, bear the lowest attendance rates. The second most proximate measure of attendance, that of the household survey, in which primary caregiver estimate their cohort girl's attendance for the past month, has the second highest

attendance rates, and the highest attendance rates were collected in the school survey which capture attendance of girls since school has started.

Given the lack of convergence in attendance rates, the headcount attendance rates are considered the baseline and midline attendance rates since the headcount attendance rates were gathered by an independent third party.

*Table 56: Measures of Girls' Attendance in Baseline and Midline*

Time frame	Survey	Baseline		Midline	
		Intervention	Comparison	Intervention	Comparison
Today	Headcount	81.4	83.7	77.1	75.5
Past month	Household	90.7	89.8	90.6	91.1
This year	School	91.9	94.5	91.7	94.3

The relationship between both the household survey and the school survey attendance rates with the headcount attendance rate is generally positive, as shown in the figure below presenting the average attendance rates of the headcount survey for a given school against the average attendance rate of the household survey and the school survey of the same school. As can be seen, in both the baseline and the midline, the attendance rates of the sampled schools from the household survey and the school survey tended to be higher than that of the attendance rates of the headcount survey. Only in the midline comparison of headcount survey attendance rates with household survey attendance rates is the correlation negative, although not significant. Supporting evidence for more accurate attendance recordkeeping by schools in the midline, the relationship of headcount survey attendance rates with school survey attendance rates is positive and significant.<sup>198</sup> The findings suggest that teachers are not only collecting more attendance records than they did in the baseline, but that those records are also more accurate. This change may be due to the project's emphasis in greater involvement of CEC's in the management of the school, and it may also be due to awareness that the attendance records are being evaluated for completeness.

<sup>198</sup> The finding is significant with a P-value of 0.039 using linear regression and excluding 6 schools in which attendance rates for the school in either the headcount survey or school survey were lower than 50 percent.

Figure 20: HH and SS Attendance Rates by HC Attendance Rate – Baseline

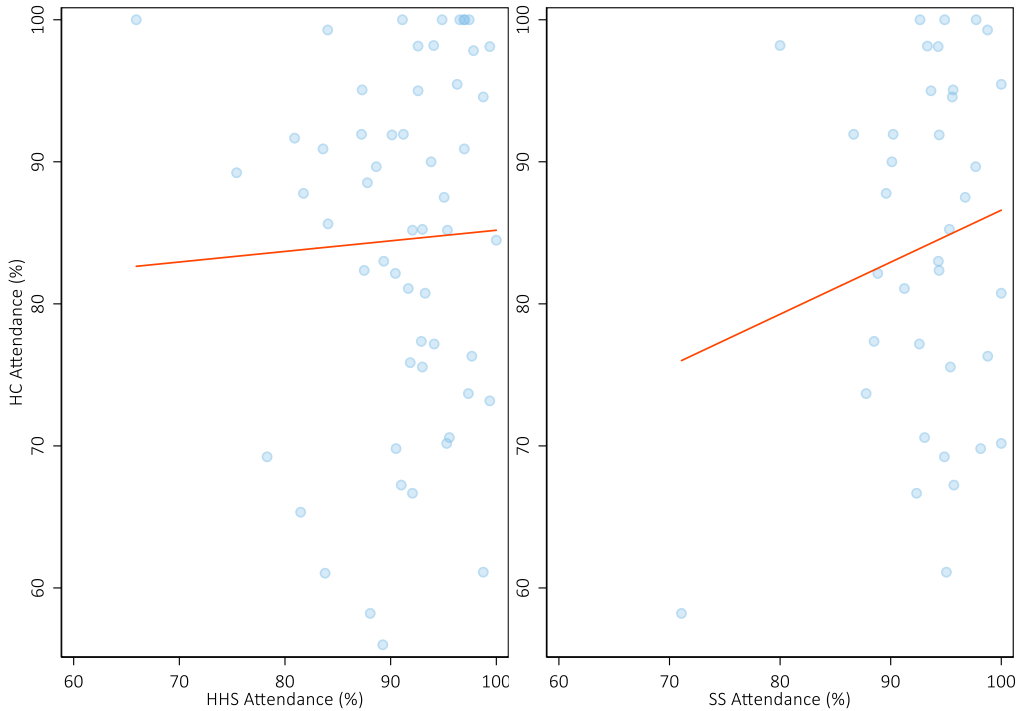
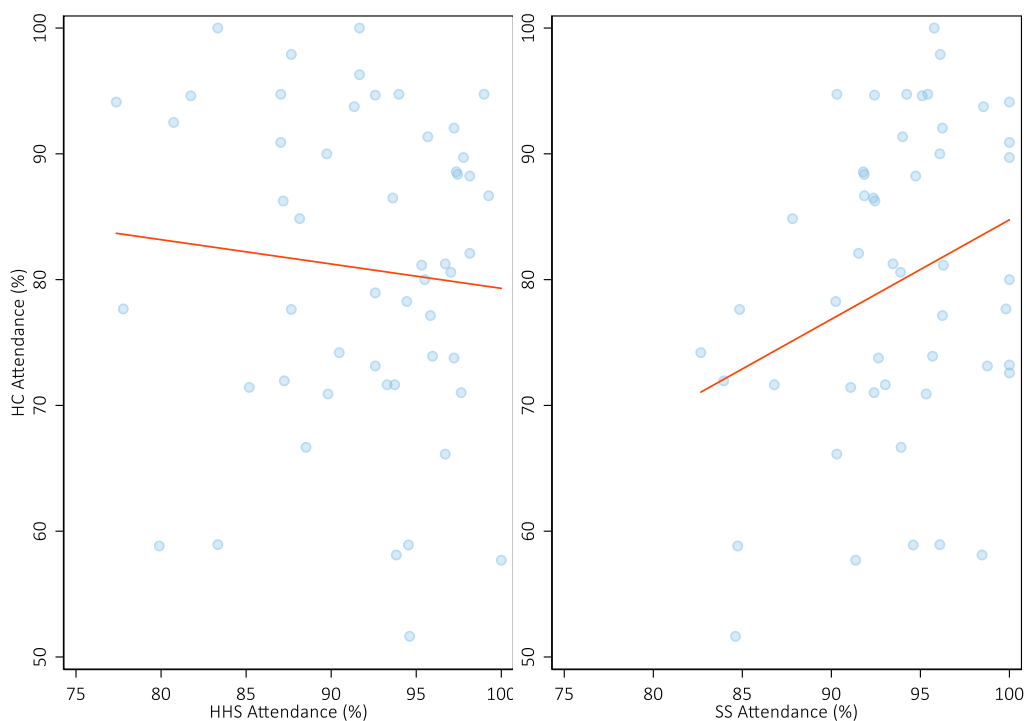


Figure 21: HH and SS Attendance Rates by HC Attendance Rate – Midline



### Progress to Attendance Targets

One of the measures of progress for the attendance intermediate outcome is the extent to which the girls in the intervention areas have higher rates of attendance in the midline evaluation as compared to the baseline evaluation. The project’s target for the attendance rates of girls in intervention schools, not including girls in ALP, by the midline evaluation point was an improvement of 5 percentage points over the baseline attendance rate. The headcount attendance rate of intervention girls in the baseline is 81.4 percent, and so the target for the midline is 86.4 percent. However, in the midline, the headcount attendance rates of intervention girls fell marginally to 77.1 percent, a decline of 4.3 percentage points and well short of the 86.4 percent goal set out for the midline evaluation point.

As mentioned in the above sections, attendance declined in the midline not only in intervention areas but also comparison areas. Indeed, the decline in attendance rates among girls in intervention areas was less than the decline experienced among girls in comparison areas. The attendance rates of girls in comparison areas declined from 83.7 percent to 75.5 percent, a difference of 8.2 percentage points. The relatively smaller decline in attendance rates by girls in intervention areas suggests that the intervention may have provided a buffer to external factors that led to decreases in attendance in both intervention and comparison areas. The absence of a statistically significant difference in difference in headcount attendance rates may at least partly be due to the fact that the intervention has only been in effect less than a year.

In the SOMGEP-T Logframe, there is one qualitative indicator for the attendance intermediate outcome: mothers’ support to adolescent girls’ attendance. The target for this indicator at the baseline was “Mothers support education and show increasing appreciation for education as a means of obtaining better jobs;

but most still prioritize domestic chores and pastoral work over attendance.” In the midline, the target for this indicator was “Mothers express support to adolescent daughters’ attendance to ASLP and formal school.” The main sources of data for the analysis of the attendance qualitative indicator were FGDs with mothers, interviews with girls with disabilities, interviews with regional education officers, and FGDs with CEC members. As was discussed above, household chores and work are still major barriers to attendance for girls at the midline. Almost every girl who has a disability that was interviewed in the qualitative portion of the research explained that she misses school only when she is sick or when her mother or parents ask her to help out at home or with a livelihoods-related task. For example, one girl explained, “I miss a day of school when I’m sick or when my mother has to go somewhere and I have to look after her shop.”<sup>199</sup> Another girl shared, “Sometimes I miss school to help my mother with herding livestock.”<sup>200</sup>

There is, however, also some data to suggest that mothers are personally taking on a higher chore burden for the explicit purpose of freeing up their daughters for school and schoolwork. One mother explained how she supports her daughters’ education: “I stopped making her help around the house with household chores, so she can prepare her lessons and go to school early, and I support her in any way I can further her education.”<sup>201</sup> It was clear in the baseline qualitative data, and is again clear in the midline data, that mothers almost unanimously appreciate the intrinsic value of education and understand the benefits associated with educating girls.

Now, it would appear that mothers increasingly appreciate that education is a means of obtaining better jobs and improving livelihoods – again, in the baseline, there was already evidence that mothers understood this link between girls’ education and jobs, but financial concerns, chores, and other factors were still often prioritized over girls’ education. In the midline, there is one major difference that has been observed: mothers appear to be supporting girls’ education *despite* the financial difficulties and personal burden it puts on them, for the explicit reason that they believe sending their girls to school will benefit the family in the future. As one CEC member explains, “Mothers put special efforts in sending their daughters to school since they have lost their livestock and would like their daughters to be able to help them and take care of them in the future.”<sup>202</sup>

The negative effects of the drought on attendance have already been discussed above; in times of extreme fragility and economic stress, many parents do clearly still keep their girls home from school to help with chores and work. However, when the financial burden is removed (i.e. when girls’ school fees are covered), a high enough number of families that are facing extreme economic stress are choosing to send their girls to school that multiple qualitative interview respondents reported the phenomenon. As one mother explains, people “had to move to the cities because of the drought and schools became free of charge. Therefore, people decided to enrol their kids at school.”<sup>203</sup> In the aggregate quantitative attendance numbers, we are not seeing evidence of these changes, but the qualitative suggests that parents are indeed making the decision to send their girls to school even during times of extreme stress. The evidence in the qualitative data suggests that the opportunity cost calculation parents are making when deciding whether or not to send their girls to school or keep them in school is beginning to shift slightly with changing circumstances and exposure to programming.

---

<sup>199</sup> IDI – GwD

<sup>200</sup> IDI – GwD

<sup>201</sup> FGD – Mothers

<sup>202</sup> FGD – CEC

<sup>203</sup> FGD – Mothers



A number of qualitative interview respondents also attributed increases in attendance at ALP schools or informal classes to programming. When asked whether attendance has changed in their community, one REO replied, “Yes, it has changed. Previously the number of girls attending school was very small, but thanks to the encouragement and payment of fees by CARE, the number has increased significantly. Girls who dropped out of school are now attending the ALP program. Old people are receiving informal education. There is now a loan association group.”<sup>204</sup> A CEC member provided further evidence that the program is encouraging parents to send their girls to ALP schools or informal classes and that even older women are encouraged to attend: “Also, there are old women who learn with the children. I used to be a teacher but today, I learn something from informal education programs.”<sup>205</sup>

### **Relationship between Corporal Punishment and Attendance**

The ToC hypothesizes that improving teaching quality will lead to higher student attendance. One measure of better teaching quality, among others such as the use of group work, gender equitable practices, and remedial support, is less frequent use of corporal punishment. In-school cohort girls were asked about the use of corporal punishment at their school. These answers were then compared with measures of the girl’s school attendance as recorded by the school survey and by the household survey.

#### *Household Survey Questions*

TQ\_7sa. How do the teachers punish students? Physical punishment?

TQ\_8s. Think about the past week at school, or the last week you were in school. In that week, did you see a teacher use physical punishment on other students?

TQ\_9s. Think about the past week at school, or the last week you were in school. In that week, did the teacher use physical punishment on you?

#### *Corporal Punishment and School Survey Attendance Rates*

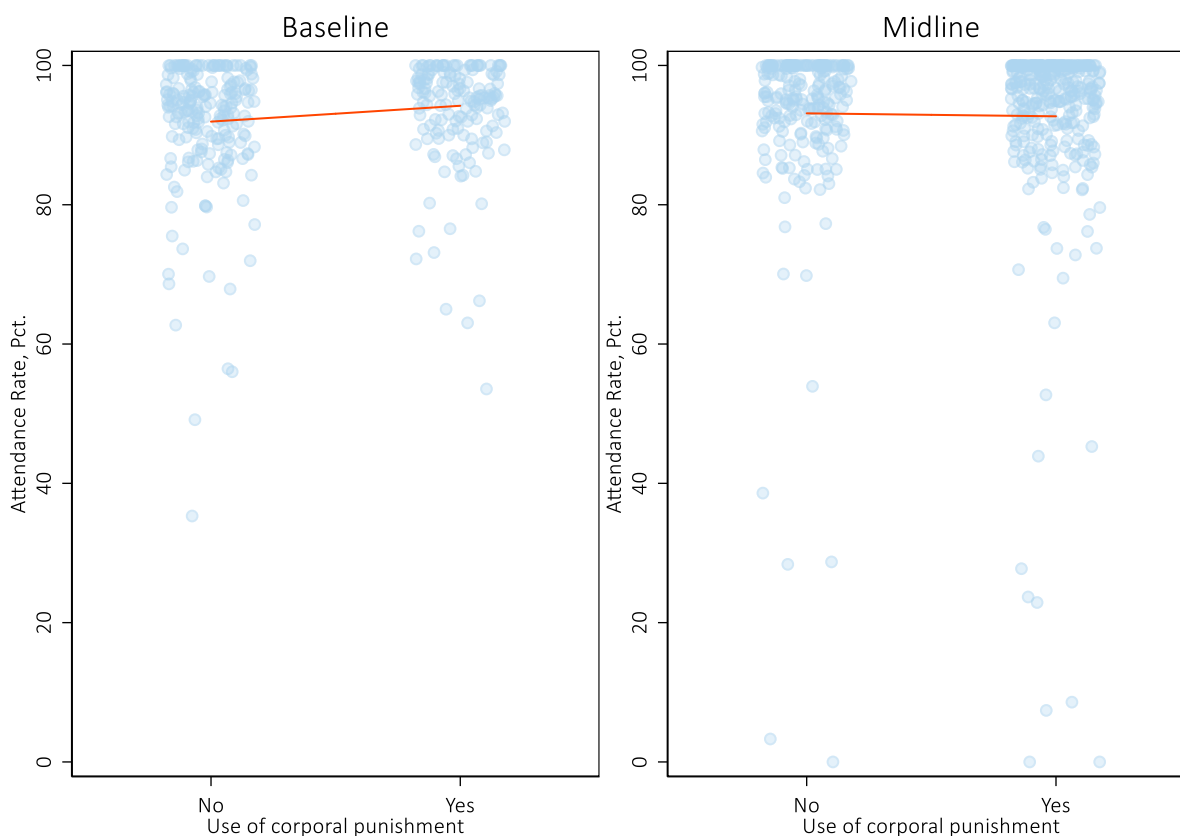
In the baseline, the cohort girls who had teachers they said used corporal punishment on students (94.2 percent) counterintuitively had higher rates of attendance than cohort girls who had teachers who did not use corporal punishment (91.9 percent), and it was a finding that was at odds with SOMGEP-T’s theory of change which seeks to decrease the use of corporal punishment. In the midline, we find that cohort girls with teachers who they say do not use physical punishment have higher attendance rates than cohort girls whose teachers do use physical punishment, albeit only marginally higher (93.1 percent vs. 92.7 percent). The figure below presents the relationship between the use of corporal punishment by the cohort girl’s teacher and the cohort girl’s school survey attendance rate.

---

<sup>204</sup> KII – REO

<sup>205</sup> FGD – CEC

Figure 22: School Survey Attendance by Teacher's Use of Corporal Punishment



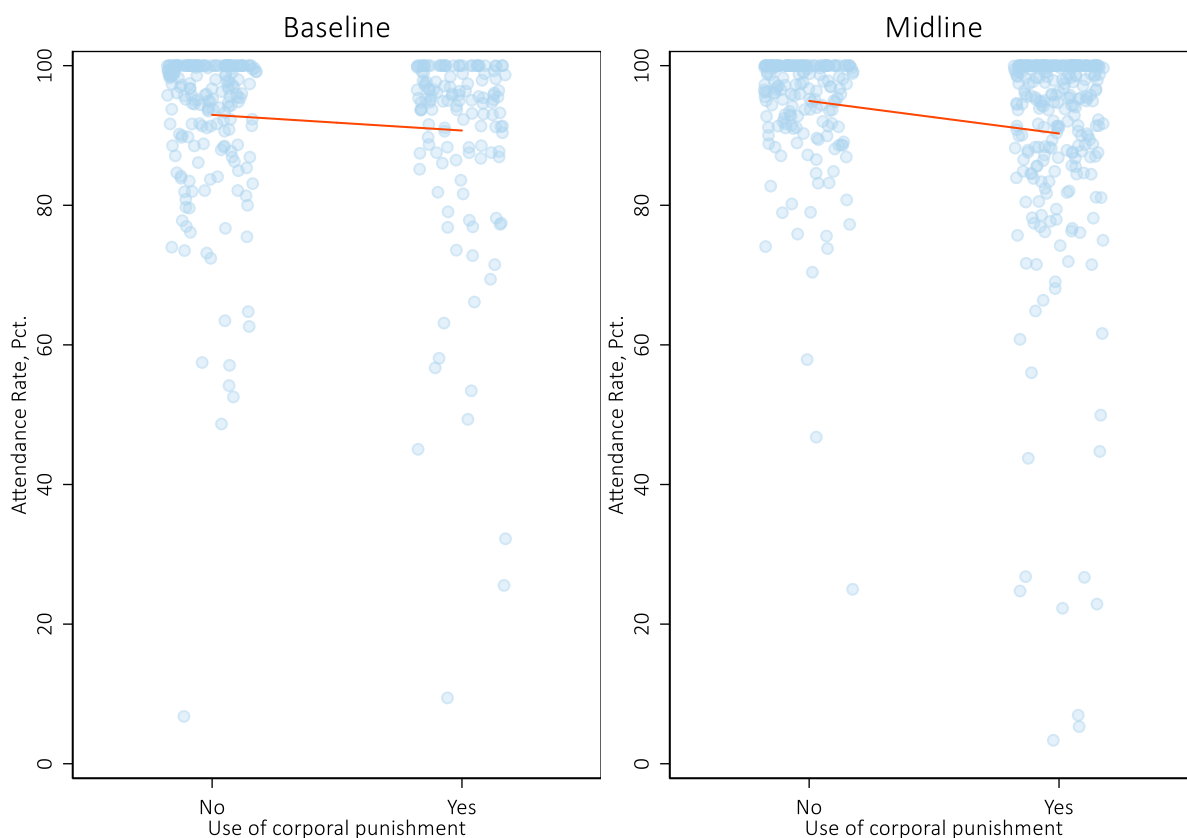
Moreover, as in the baseline, in the midline there is not a significant relationship observed between school survey attendance rates and the girls' responses about the use of corporal punishment by the teacher on other students or the girls themselves during the week prior to the survey team's visit to their school.

#### *Corporal Punishment and Household Survey Attendance Rates*

The relationship between household survey attendance rates and the use of corporal punishment is somewhat less ambiguous. As shown in the figure below, in the baseline, girls who said that their teacher used corporal punishment had marginally lower attendance rates (90.3 percent) than girls with teachers who did not use corporal punishment (94.5 percent). The pattern observed at the midline is similar but the difference is now statistically significant: girls with teachers who used corporal punishment in their classroom had significantly lower attendance rates than girls who did not (90.3 percent vs 94.9 percent).<sup>206</sup>

<sup>206</sup> The finding is significant with a P-value of 0.003 using cluster robust linear regression.

Figure 23: Household Survey Attendance by Teacher's Use of Corporal Punishment



No relationship is observed between the household survey attendance rate and whether the girl said the teacher used physical punishment in the past week or whether the cohort girl said the teacher used physical punishment on her in the past week. This non-finding may in part be due to the short time frame of those questions and the relative infrequency with which teachers resort to physical punishment. In the midline, only one percent of girls in the baseline said that their teacher uses physical punishment almost every day, and only three percent of girls answered the same.

The results from the qualitative interviews and discussions suggest that many teachers consider there to be a link between corporal punishment and either poor attendance or dropping out. Some take the matter of punishment very seriously – for example, one teacher explained, “If we beat up a student who was struggling but making efforts we may destroy his/her future. That will become a problem. The student will drop out of school as a result.”<sup>207</sup> This appears to be a direct result of exposure to training that encourages other punishment techniques: “If you are a teacher, you cannot beat the student if the student did not understand something. It may have happened when we were students, but the situation is different now. Teachers are receiving lots of training about how students understand something.”<sup>208</sup>

<sup>207</sup> FGD – Teachers

<sup>208</sup> FGD – Teachers

However, it would appear as though this training has focused specifically on the effects of corporal punishment on girls, not boys. There is likely some social desirability bias at play in the qualitative results, but the responses of multiple respondents across different groups suggest that teachers reserve harsher punishments for boys. This was stated explicitly by teachers: “If a girl commits any mistake, we do not treat her like a boy; girls are vulnerable, so if we treat them like boys, they may drop out because of that. So in the class we do not talk about the mistake she committed but we send her outside and we solve her problem privately to keep her emotions.”<sup>209</sup> It was also stated explicitly by mothers, and others: “In terms of the education they receive, it is the same. But, in terms of the discipline and punishment they receive, it is not the same. Boys will receive harder punishments, whereas with girls, it is likely their parents will be informed if they have misbehaved.”<sup>210</sup>

Although it is a positive development that teachers are taking the relationship between corporal punishment and attendance seriously, the intense focus on the effects on girls appears to have had two observable, potentially negative consequences: (1) Girls and boys are not being punished equally, and (2) Girls are considered more fragile than boys. Special treatment of girls in this sense could in fact be inadvertently reinforcing negative social norms around femininity, as well as causing resentment in the classroom, although there has been no evidence of the latter. In the baseline, during which FGDs were conducted with both boys and girls, many students appeared to feel that harsher punishments are fair and reserved only for problem students. Many students and teachers also expressed that boys tend to misbehave more than girls. However, it was not possible for the purposes of this report to gauge whether these attitudes have changed since the baseline, as the girls’ and boys’ FGDs were not conducted in the midline. The effects of programming on boys should be examined in more depth in the future to fully understand the program’s intended and unintended impact in this regard.

### Girls’ Characteristics Analysis of Attendance

This section presents the attendance rates, based on caretaker’s reporting in the household survey, for various subpopulations of interest, defined by geography, disability, economic status, and other girl- and household-level characteristics. We also present attendance rates for girls who face barriers to educational attainment, including barriers related to school infrastructure, school resources, teaching quality, gender equality, and other general barriers.

The subgroup results are presented in the table below. Attendance rates are reported for each subgroup for the baseline sample in the second column and the midline sample in the third column. In the fourth column, we report the change in average attendance since the baseline shown as a percentage increase or decrease. Finally, the last column reports the number of girls in the midline transition sample who fall into the subgroup.

*Table 57: Attendance of key subgroups*

	Baseline average attendance (aggregate)	Midline average attendance (aggregate)	Change in average attendance since baseline	Number of observations for subgroup
<b>Characteristics:</b>				
All in-school girls	90.6	91.3	0.7	601

<sup>209</sup> FGD – Teachers

<sup>210</sup> FGD – Mothers

Living without both parents	91.7	90.7	-1.1	59
Mother tongue different to LOI	94.4	94.7*	0.3	31
<b>Disability</b>				
Vision impairment	93	88.3	-4.7	3
Hearing impairment	0	0	0	0
Mobility impairment	0	100.0*	0	1
Cognitive impairment	0	0	0	0
Self-care impairment	0	0	0	0
Communication impairment	100	100	0	1
Mental health impairment	92.5	89.4	-3.1	62
Anxious	93.5	90.6	-2.9	53
Depressed	92.3	89.4	-2.9	39
Any disability	92.5	89.4	-3.1	65
<b>HOH and Carer Characteristics</b>				
HOH no wage-earning occupation	90.7	92.8	2.1	305
HOH no education	91.1	91.5	0.5	290
HOH female	92	92.7	0.8	274
HOH pastoralist	87.4	83.2*	-4.2	44
Carer no education	90.1	91.2	1.1	322
<b>Household Assets</b>				
Owns camels	88.8	90.2	1.4	65
Owns medium-sized livestock	89.9	90.3*	0.4	414
Owns small livestock	89.7	81.9*	-7.8	32
Owns mobile phone	90.5	91	0.5	539
Access to water reservoir/storage	91	91.4	0.4	366
Owns land	91.3	90.9	-0.3	426
<b>Poverty</b>				
House is informal/temporary structure	89.1	91.1	2	35
Gone to sleep hungry most days	83.8	81.8	-2	24
Gone without enough clean water most days	87.6	90.1	2.5	70
Gone without medicines or medical intervention most days	89.2	91	1.9	129
Gone without cash income most days	87.8	92.3	4.5	141
Migration	98.7	96.2	-2.5	18
Displaces or moved in past 12 months	78.3	80	1.7	4
Household migrates seasonally	86.6	77.6	-9.1	9
<b>Other</b>				
High chore burden (whole day spent on chores)	92.1	89.2	-2.9	51
Married	82.4	79.5*	-2.9	5

Mother, under 16	100	86.8	-13.2	3
------------------	-----	------	-------	---

\*Note, an asterisk indicates results that are statistically significant at the 95% confidence level (or higher) in a regression with cluster-robust standard errors.

The sample sizes for many of the disability-related subgroups are too small to draw meaningful conclusions about attendance, but as has been noted in previous sections, the sample does include high enough numbers of girls with mental health issues to assess the impact of mental health on attendance rates. Although disability is not a significant predictor of attendance rates, girls who exhibit mental health issues have slightly lower attendance rates than the rest of the sample.

In keeping with the analysis of learning outcomes, many proxies of pastoralism are also strong predictors of lower than average attendance at the midline. Having a pastoralist head of household, as well as ownership of medium-sized livestock and of small livestock are all significant predictors of lower than average attendance rates. It is also worth noting that more general indicators of poverty are not predictive of attendance rates, suggesting that the project may have mitigated some of the challenges that poorer families were facing at the baseline. And, while the sample size of married girls is small, marriage is a significant predictor of lower attendance rates, suggesting that these girls are at high risk of dropping out of school.

In the learning outcome section, the finding that pastoralism is a predictor of lower than average learning outcomes was explained in part by the effects of recent droughts, which appear to have been experienced most acutely by pastoralist families. In terms of attendance, we know that, in light of the recent droughts, pastoralist households are particularly economically distressed, as many have lost all their livestock. The qualitative data suggests that these households are therefore more likely to keep their children home from school as a source of labour. As one mother explains, “The quality of education decreased during the drought. Parents sent their children to collect jerry cans instead of attending school.”<sup>211</sup> Lower attendance is predictor of lower learning outcomes, which also helps to explain why pastoralism was a predictor of lower learning outcomes.

However, as was noted in the transition section above, overall enrolment has increased. Again, the qualitative data holds some insights that could help to explain this finding. It appears that the drought has in some cases had the counterintuitive effect of leading more pastoralist families to enrol their children. A number of qualitative interview participants in the midline outlined ways in which drought has increased access to schooling for pastoralist children. Multiple participants explained that families who may not have sent their girls to school otherwise are now invested in their girls’ education. One teacher explains, “The drought actually helped encourage students. When the droughts happened, all of the pastoralists’ livestock died so they moved to the villages and were able to send their children to school. Teachers adapted to this and were able to teach more students.”<sup>212</sup> A CEC member explains, “Mothers put special efforts in sending their daughters to school since they have lost their livestock and would like their daughters to be able to help them and take care of them in the future.”<sup>213</sup> A mother from a different community expressed a similar sentiment: “Most of the people that came here were pastoralists whose livestock died due to the drought. Now that they are here they have no other option but enrolling their children at school.”<sup>214</sup> In the same interview, another mother added that people “had to move to the cities

<sup>211</sup> FGD – Mothers

<sup>212</sup> FGDs – Teachers

<sup>213</sup> FGDs – CECs

<sup>214</sup> FGDs – Mothers

because of the drought and schools became free of charge. Therefore people decided to enroll their kids at school.”<sup>215</sup>

These two findings – that enrolment has increased and attendance has not – appear at first glance to be contradictory. However, the qualitative findings outlined above suggest that pastoralist families settling down as a result of the drought may be driving overall enrolment rates. This would also explain why attendance rates have not followed suit – pastoralist families, who are as a result of the drought in extreme economic distress, are also likely to keep their children home from school at least some of the time to help with labour or stay home while one or both parents seeks supplementary income. We may simply be seeing an increase in enrolment among a subset of the population that tends to attend school less.

### Barriers Analysis of Attendance

The next table presents the attendance rates for girls facing a number of barriers to educational attainment. This table follows the same structure as the table above, presenting first baseline attendance rates, then midline rates, the change in average attendance from baseline to midline, and the number of observations for each barrier.

Table 58: Attendance of key barriers

	Baseline average attendance (aggregate)	Midline average attendance (aggregate)	Change in average attendance since baseline	Number of observations for barrier
<b>Barriers:</b>				
All in-school girls	90.6	91.3	0.7	601
<b>School Infrastructure</b>				
Difficult to move around school	91.9	91.4	-0.5	123
Doesn't use drinking water facilities	88.8	92.3	3.5	110
Doesn't use toilet at school	93	91.4	-1.6	160
Doesn't use areas where children play/socialise	90	91.4	1.4	256
<b>School Resources</b>				
No computers at school	90.4	91.4	1	566
School does not have learning materials	91.5	91.7	0.2	165
Not enough seats for children at school	89.5	83.8	-5.8	97
<b>Teaching Quality</b>				
Disagrees teachers make them feel welcome	92.7	90.3	-2.3	26
Agrees that they are afraid of teacher	90.5	91.4	1	468
Agrees teachers treat boys and girls differently in the classroom	89.9	91.7	1.8	242
Agrees teacher is often absent from class	88.6	92.4	3.8	132
Teacher punishes students who get things wrong	90.4	91.2	0.8	512
Teacher uses corporal punishment	89	90.3	1.3	190

<sup>215</sup> FGDs – Mothers



Carer says principal performance is poor	78.8	80.1*	1.3	17
Carer says teaching at school is poor	85.3	84.5	-0.8	16
<b>Gender Equity</b>				
Teacher targets questions by gender	90.2	92.4	2.2	62
Teacher targets difficulty of questions by gender	92	92.4	0.5	87
<b>Other Barriers</b>				
Agrees she has no choice in schooling decisions	90.5	91.3	0.7	536
Over 30-minute travel time to school	94.2	90.9	-3.3	39
Feels unsafe on way to school	92.8	93.2	0.4	11
Feels unsafe at school	95.6	90.6	-5	8

\*Note, an asterisk indicates results that are statistically significant at the 95% confidence level (or higher) in a regression with cluster-robust standard errors.

Most barriers were not significant predictors of caretaker-estimated attendance rates. The only barriers that were strongly predictive of lower attendance rates were those related to teaching quality. Poor principal performance was a statistically significant predictor of lower attendance. That is, girls who went to schools with principals who were rated poorly by caretakers had on average a significantly lower attendance rate, 80.1 percent, than girls whose caretakers did not, 91.0 percent. The qualitative data suggests that if parents think that a school is poorly run, then they are less likely to send their girl to school, which helps to explain this quantitative finding. For example, one mother explains that she pulled her daughter out of school because of the poor quality of schooling: “When my daughter was studying there, I found out after 2 years that she did not understand or learn much. I followed up on it and found out that she was not being taught anything at school. So I took her out of school and nobody came to me to follow up on why I had taken her out of school.”<sup>216</sup> While not statistically significant, poor teaching<sup>217</sup> and having insufficient seats for children at school<sup>218</sup> also predicts attendance that is substantially below average.

The SOMGEP-T project seeks to improve the intermediate outcome of attendance by enhancing the capacities of community education committees, equipping girls with school supplies and paying for school fees, engaging with community-level stakeholders, supporting girls’ education forums and boys’ education forums, VSLAs, and working with the Ministry of Education to enhance their capacity to deliver formal and informal educational services. The project’s design can be described as GESI Transformative in that it “actively seeks to transform inequalities in the long term for all children despite gender, disability or other characteristic.”<sup>219</sup> The project not only aims for short-term improvements in attendance through bursaries, school equipment, and GEFs, but it also aims for a systematic change of the society in which the Somali girls live by improving the households’ ability to afford to continue sending their children to school, school governance mechanisms, addressing community attitudes about the education of girls or other marginalized groups, and the MoE’s capacity to deliver educational services.

These project interventions have not yet had a measurable impact on attendance in intervention areas relative to comparison areas, likely in part because the intervention has been underway for less than a

<sup>216</sup> FGD – Mothers

<sup>217</sup> The P-value is 0.094 for the regression of caretaker-estimated attendance on the poor teaching quality variable.

<sup>218</sup> The P-value is 0.119 for the regression of caretaker-estimated attendance on the insufficient seats variable.

<sup>219</sup> Guidance for EE on GESI Midline Dec 2018, *Girls’ Education Challenge*, December 2018.

year. Nevertheless, some progress as well as areas to focus on have already been identified. Headcount attendance rate of girls in intervention areas was 82.0 percent in the midline an only marginal improvement from 81.9 percent in the baseline, but one which is greater than the decline of attendance rates among girls in comparison areas from 84.0 percent to 81.0 percent. Attendance recordkeeping has become more common. In the baseline, attendance records were found for only 48.9 percent of girls but this proportion rose to 70.3 percent in the midline. Furthermore, based on comparisons of the attendance records in the school survey and the headcount data, the attendance records appear also to be more accurate in the midline than they were in the baseline. Finally, we find that pastoralism, being married, and poor principal performance as rated by the caregiver are predictors of lower levels of attendance. This finding confirms the ToC change that marginalised girls include those from pastoral families and that early marriage and poor school management are barriers to education.

## 6.2. School governance and management

Improved school governance and management is the second SOMGEP-T intermediate outcome. It is not only important for the sustainability of the project, but is also an essential outcome for improvements in students' learning and transition. The purpose of this section is to assess the degree to which SOMGEP-T has realized its intended outcomes with respect to school governance and management affairs and determine improvements needed to achieve its goals in the project's remaining years. Thus, Difference in Difference (DiD) is employed to compare the change over time for the school governance and management in the intervention and comparison groups. DiD controls for differences between the groups over time that were not a result of the program.

The analysis utilizes the same key indicators established in the quality of school governance baseline as well as some new indicators.<sup>220</sup> Unlike the baseline study, no teachers' survey took place for the midline study, and the indicators are based on two distinct surveys: a survey of head teachers and the household survey, completed by primary caregivers. Therefore, the results presented below fall into two main categories: the first category examines the extent of Community Education Committees' (CECs) establishment and performance from the head teachers' perspectives, while the second category assesses CECs and school management by asking the primary caregivers to rate how well the school is managed and how head teachers performed. The data is disaggregated by zone.

### Community Education Committees' Assessment from Head Teachers' Perspectives

In the absence of strong and effective central education authorities, CECs play an important role in the management of schools in Somalia. The committees' members consist of local volunteers including parents, religious leaders, head teachers, and members of women's and youth groups.<sup>221</sup> The CECs hold various responsibilities, such as acting as a liaison between the school and the community, monitoring school and students' performance, overseeing school policies and teaching quality, and encouraging students' enrolment. By providing financial and non-financial support to schools, CECs also help to ensure that SOMGEP-T's achievements are sustained after the end of the project.

In addition to assessing the presence of CEC in schools since baseline, the following analysis explores the extent to which the CECs were engaged in various activities, such as school monitoring,

---

<sup>220</sup> Due to lack of baseline for the new indicators, we have only reported the descriptive analysis of the midline data. This includes indicators measuring CEC's existence in schools, frequency of meetings, membership, overall performance, and monitoring visits and activities, and scholarships.

<sup>221</sup> James H. Williams & William C. Cummings (2015) Education from the Bottom Up: UNICEF's Education Programme in Somalia, *International Peacekeeping*, 22:4, 419-434, DOI: [10.1080/13533312.2015.1059284](https://doi.org/10.1080/13533312.2015.1059284)

communication of plans to parents and community members, and the provision of financial and in-kind support to the school. Table 59 shows regression and difference-in-differences results across all analytical variables.

The first step towards establishing a well-governed school is to have an established CEC where the members meet regularly before they get involved in more complicated aspects of school management, such as providing financial support to school. When asked to report the presence of a CEC in the midline, 96.8 percent of the head teachers in the entire sample said that they have a CEC in their schools. When data were disaggregated by intervention and zone in the midline, the presence of CECs did not significantly differ across groups.

However, it is important to mention that the presence of CECs indicator is established based on head teachers' perceptions and does not therefore provide a definitive measure of whether they actually exist. Quantitative and qualitative evidence indicates that some head teachers are not probably aware of a CEC in the school, while others might wrongly assume that CECs do not exist. Qualitative data shows that in some locations, CEC members describe their active involvement in schools and in their communities, and this involvement is confirmed by the responses of teachers and mothers. In other locations, CEC members describe being involved in school activities, such as monitoring, but teachers appear to be completely unaware of the existence of a CEC. For example, in one community, a CEC member shares that the CEC sends members to monitor "how teachers and students work together, whether there is conflict, and whether there is food for the students,"<sup>222</sup> but teachers from the same school question whether there is a CEC at all, with one explaining, "before there was a CEC. Nowadays, the CEC has been destroyed. The community must come together and decide to create a CEC."<sup>223</sup> Mothers from the same community, however, confirm that the CEC raises awareness on girls' education, raises the financial support for girls to go to school, and monitors teachers and students. This specific case may simply be an example of poor coordination between CEC members and teachers. Therefore, the presence of CEC indicator is believed to be a more accurate measurement of CECs' establishment as well as their activity level rather than their establishment only.

*Table 59: Project impact (difference-in-differences) on school management and governance indicators*

Indicators	Baseline/ Intervention	Midline/ Intervention	Difference- in- Differences
School have a functioning CEC (n=63)	-	0.03	-
Overall performance of the CEC in managing school (n=61)	-	0.15	
Frequency of CEC meetings (n=61)	-	-0.23	-
Number of CEC members (n=61)	-	0.03	-
CEC member monitored the school in past year (n=63)	1.61**	1.2	-0.17

<sup>222</sup> FGD – CECs

<sup>223</sup> FGD – CECs

Frequency of CEC's monitoring (n=43)		-	-0.83	-
During the last visit by a CEC member, what did they monitor?	Teacher's attendance	1.6	-0.69	-2.3
	Facilities (n=49)	0.81	0.43	-0.38
	Teaching quality (n=49)	-0.18	-1.59*	-1.4
	Students' attendance (n=49)	1.23	0.49	-0.74
	Student retention (n=49)	0.37	-0.84	-1.21
Does the school have a school/CEC management plan? (n=63)		1.65*	0.58	-1.07
Over the last six months, what did CEC do?	Monitored student attendance (n=63)	-	-0.19	-
	Followed up on students' dropouts (n=63)	-	-0.2	-
	Monitored Teachers' attendance (n=63)	-	-0.24	-
	Took actions against teachers with irregular attendance (n=63)	-	0.59	-
	Raised funds (n=63)	-	0.97	-
	Reinforced use of non-violent discipline (n=63)	-	0.71	-
	Addressed child protection issues (n=63)	-	0.77	-
	Promoted enrolment (n=63)	-	0.9	-
Total amount contributed to teachers' salary by CEC (n=53)		-38.51	-45.3	-83.9
Teachers received incentives to stay in school (n=63)		0.79	-0.5	0.02

In the midline, a large majority of the head teachers rated the performance of CECs as either “very good” (36 percent) and “somewhat good” (52.4 percent). Yet, in order to gain a better assessment of the CECs’ performance, it is not only important to understand how they accomplish their responsibilities and tasks, but also how they perform as an organization, such as the frequency of the meetings they hold and the number of members a CEC is comprised of. According to the school survey, the most frequent number of times that CECs meet is once per month (44.2 percent), followed by once every two week (29.5 percent) and once every two months (14.8 percent). Additionally, more than 90 percent of the CECs have 5 to 7 members, including the head teachers. The differences between the groups in the midline were significant neither in terms of how often they hold a meeting nor how many members a CEC acquires.

Another indicator measuring the activity level of CECs is whether a member of the school CEC monitored the school in the last year, and if so, how many times. According to the head teachers, 77.8 percent of CECs visited schools at least once in the past year, but we do not observe significant difference-in-differences among the intervention and comparison group nor across locations. The negative sign of difference-in-differences coefficients implies that the percentage of CECs in the intervention group who have monitored schools in the past year has increased less than their counterparts in the comparison group. On average, CECs completed 13 school visits in the past year.

In the midline, the intervention group also did not significantly differ compared to the comparison group with respect to how many times they have visited schools. Again, a significant difference-in-differences was found neither between the intervention groups nor across locations on whether the CECs monitored teachers’ attendance, facilities, teaching quality, or students’ attendance and drop out during their last visit. The negative difference-in-differences figures may either suggest that the CECs have substantially reduced the frequency of the times they carry out those specific activities or the progress in the intervention group was smaller than the comparison group.

Furthermore, having a school management plan is the second most fundamental aspect of a well-organized school governance after the establishment of a CEC. In the midline, less than half of the CECs (46 percent) were reported to have a management plan. As Table 59 indicates, no significant difference-in-differences were found when comparing changes in the intervention group with changes in the comparison group with respect to having a school management plan. Again, the negative sign associated with the difference-in-differences coefficient is an indication of smaller increase in the number of schools in the intervention group having a CEC plan compared to the comparison group.

The results from the FGDs with CEC members, mothers, and teachers suggest that the current capacity of CECs varies widely depending on the area, but that CECs in almost all areas are following up on cases of dropout, monitoring schools, carrying out awareness raising activities, and fundraising when possible. It is less clear from the results what stage a given CEC is at in developing formal school management plans. When asked whether there is a school management plan at the school in their community, most CEC groups reported that there is a plan, but their responses suggest that the concept of a school management plan is not necessarily uniform across schools. In some areas, CECs described formal school management plans, such as five-year plans, whereas in others, plans appear to be informal and include ideas discussed during meetings between CEC members, between CEC members and teachers, or between CEC members and parents. For example, when asked if the school in their area has a school management plan, one CEC member responded with the following: “Recently, there was a committee.

We created a committee to encourage boys who dropped out to return to school.”<sup>224</sup> Another responded, “Yes, we have a plan. We always have meetings on Thursdays between teachers and the principal of the school.”<sup>225</sup>

The head teachers reported that their school management plans include policies about monitoring the school (student attendance, teacher attendance and teacher practices) (82.7 percent), child protection policies (69 percent), encouraging school enrolment (82.8 percent), and plans to follow up with dropouts (79.3 percent).

Moreover, qualitative data suggests that there are some tensions between the head teachers and CEC members. Teachers mainly appear to take issue with CEC members’ lack of experience and exposure to training. As one teacher explains, “there is a big misunderstanding between the CEC and the teachers because most of the CEC members are not educated and they do not understand their role. The majority of challenges are caused by them because of their role [because they do not understand their role]. We teachers have complained to the community leader, who was asked to educate the CEC members about their role. As a consequence, two members were replaced. The new members are better educated, although there are still some misunderstandings.” At another school, a teacher echoes this sentiment: “I think they do not understand the system of the work and they need to receive more training.”<sup>226</sup> In yet another school, one teacher explains that the challenges facing the CEC include “a lack of training for the CEC and a lack of relevant knowledge to succeed in their work”<sup>227</sup> and another explains, “communities may disagree with the principles of the committee. For example, parents might disagree with the committee’s decision to send a student who came late back home.”<sup>228</sup>

These same tensions do not appear to exist in other locations. One CEC member explains, “for example, we used to collect the funds of the schools separate in the past years, but this year all the funds for the three schools have been put into one pool and they are managed from the same point.”<sup>229</sup> Another clarifies, “yes, the funds for all schools have been put into one place. We established a committee who will manage the funds and provide the teachers with their salary that they used to receive before. That plan has covered the needs of the teachers whereas previously they were not receiving regular salaries.”<sup>230</sup> In one school, there are regular meetings between teachers and CEC members, CEC members are actively involved in school events, and again, CEC members played an active role in raising money for teachers’ salaries. One teacher shares his views on CEC members in his community: “the CEC deserves a Noble Peace Prize because most of them are mothers who have children, but when I saw how they are interested in the education, I wish to give awards to motivate them. They are available all the time and their goal is to monitor and observe teachers and students. They check the challenges and they support the school.”<sup>231</sup>

The major differences between these schools and the aforementioned schools appear to be: 1) the level of prioritization of teachers’ needs and concerns, and 2) the level of communication between CEC

---

<sup>224</sup> FGD – CECs

<sup>225</sup> FGD – CECs

<sup>226</sup> FGD – CECs

<sup>227</sup> FGD – CECs

<sup>228</sup> FGD – CECs

<sup>229</sup> FGD – CECs

<sup>230</sup> FGD – CECs

<sup>231</sup> FGD – CECs

members and teachers. One group of teachers either explicitly stated that the CEC members do not listen to their needs or implied that CEC members do not understand the education system or the challenges teachers face. Despite this, CEC members are in somewhat of a position of power over teachers in that they are involved in monitoring activities and make suggestions to administration on how to improve the school.

The project aimed to increase CEC's monitoring student attendance and student retention by 70 percent from the baseline to midline, however, the project did not seem to have reached its goal entirely. At the aggregate level, slightly more than half of the CECs (52.3 percent) reported monitoring student attendance while 58.7 percent of CECs follow up with or contact dropouts. When data is disaggregated by intervention, it appears that the comparison group have insignificantly more CECs that monitor student attendance (54.8 percent) and follow up on dropouts (61.2 percent) compared to the intervention group (50 percent and 56.2 percent, respectively).

In the project logframe, there is one qualitative indicator on school governance: CECs' perceptions on the importance of the retention of marginalised sub-groups, such as pastoralists. In the baseline, the target for this indicator was "CECs are following up on cases of dropout, but still consider that it is normal for pastoralists to miss school." In the midline, the target for this indicator was "CECs include retention of marginalised groups in school improvement plans." FGDs with CEC members, FGDs with teachers, and IDIs with girls with disabilities were used to inform the analysis of this indicator in the midline.

Despite a lack of consensus over what constitutes a school management plan, it is clear from the FGDs that CECs in most communities are making a concerted effort to follow up with students who have dropped out or who have not yet enrolled in school due to financial difficulties. Money raised through fundraising appears to most commonly be used to cover school fees for marginalised children, teacher salaries, and school improvement projects. As one CEC member explains, "Yes, due to the heavy impacts of the drought some students could not afford to pay school fees. We have managed to bring back students who have left with the help of Allah and now they are all studying again."<sup>232</sup> Another explains, "Yes – we raised money two times, and the reasons were low salaries for the teachers, which came from CEC, and the other time was when the school was neglected and the CEC had only 7 members who provided repairs for the school, such as chairs, boards, and walls. The CEC raised money to help some teachers with no salary to receive it."<sup>233</sup>

The one group CECs and schools do not appear to be focusing an adequate amount of attention on is children with disabilities. In some cases, this appears to be due to the fact that there are not disabled children in the community. One teacher explains that "the disabled people are very few in this village – the only few disabled are adults and in the afternoon we teach them for free."<sup>234</sup> However, in other areas, teachers admit that there has been more of a focus on students who cannot afford school fees or orphans than on disabled students: "I remember a few students that the school used to help -- they were poor students and orphans, but we did not help disabled students."<sup>235</sup> In other cases, teachers explained that there is a general lack of knowledge or financial ability to assist children with disabilities: "No, our school has not taken action to help children with disabilities because we do not know how to start such a

---

<sup>232</sup> FGD – CECs

<sup>233</sup> FGD – CEC

<sup>234</sup> FGD – Teachers

<sup>235</sup> FGD – Teachers



program. In addition, we do not have the required equipment or anyone who has technical expertise on children with disabilities.”<sup>236</sup>

This finding is particularly troubling, considering negative attitudes/behaviours toward children with disabilities appears to be discouraging them from attending school: “There was a blind student who used to attend school but her mother told her not to come anymore fearing the children would further damage her eyes.”<sup>237</sup> Additionally, girls with disabilities do not always feel able to participate in school activities. As one girl explains, “Yes, there are so many things that I would like to take part in but I do not have a chance to participate. For example, this school has a CEC, and sometimes that CEC arranges games for the students like playing or having concerts and I cannot do those activities because I am sick.”<sup>238</sup> There is a clear gap in support for children with disabilities that CEC members could help fill with specific training on disability-sensitivity.

Financial constraints were among the second most frequently cited challenges in the teacher and CEC FGDs. Lack of financial resources impacting CECs’ performance and schools’ sustainability is particularly noticeable because the community members were less capable of making contributions to school’s running costs due to extensive drought experienced during the past year in the area. In one FGD, the CEC members were asked whether they were able to raise any money for the school, and one of the CEC respondents stated, “We did not collect financial contributions this year to this school because people they did not have finance and there are droughts.”<sup>239</sup> Less financial contributions translates into lower capability to pay teachers’ salaries and other school expenses and support girls from low-income families.

Data from the school survey indicates that less than 1 in every 5 teachers (19 percent) received a salary from the CECs, with teachers receiving an average amount of 347.2 USD. Female teachers constitute a small portion of the teachers (8 percent) who benefited in up to 20 percent of their salaries from in cash or in-kind contributions made by CECs, whereas their male counterparts (17 percent) were paid up to 32 percent of their salaries on average. There is no significant difference in the total amount contributed by CECs to teachers’ salaries, which could be attributed to the SOMGEPT program.

When asked if teachers were provided any incentives by any groups to encourage them stay at school, only 30 percent of respondents confirmed receiving incentives in the midline. The DiD’s results did not show any significant change in the number of teachers receiving incentives in the intervention group compared to the comparison group since the baseline. Most of this type of support was provided by the Ministry of Education (42.1 percent) followed by other community groups (10.5 percent) and the CECs (5.26 percent). Besides teachers, students were also assisted financially by CECs to attend school. The school survey’s midline data shows that 58.7 percent of students received scholarships overall, of which only 18.9 percent were provided by a CEC.

### **Community Perceptions of School Management**

Besides head teachers who deal with school management in their day-to-day work, the students’ primary caregivers’ perceptions could also be an important indicator reflecting how well a school is managed. As part of the household survey, the caregivers of in-school girls were asked to assess the general

---

<sup>236</sup> FGD – Teachers

<sup>237</sup> FGD – Teachers

<sup>238</sup> IDI – GwD

<sup>239</sup> FGD – CEC

management of their children’s school. Overall, 75.6 percent of mothers reported having a CEC in school. However, like with the head teachers, there seem to be some discrepancies in parents’ responses regarding CEC presence within the same school in the household survey. For example, of the 63 schools sampled in the midline, less than 30 percent of parents provided consistent responses about the existence of CECs compared to their peers within the same school. Most of the mothers (91 percent) gave positive feedback about the management of their children’s schools, with 72.1 percent rating it as “extremely well managed” and 18.8 percent as “well managed.” A similar percentage of parents also believe that the management of schools has “improved” (79.8 percent) or “stayed the same” (13.8 percent) as compared to 12 months ago. The difference-in-differences regression analysis shows that the caregivers in the intervention group not only tend to have more CECs in their children’s schools, but also tend to be more satisfied with school management, and hold more positive attitude towards the changes they have seen in school management compared to the last 12 months. However, these differences were significant neither between the intervention and comparison group nor across other subgroups such as the parents of girls with disabilities, girls 12 and over, and across regions. The table below shows the difference-in-differences results across all analytical variables.

*Table 60: Project impact (difference-in-differences) on household measures of school management*

Indicators		Difference-in-Differences	Any Disability	Age >=12
Does the school have a CEC? (n=764)		0.091	0.015	0.05
How well is the school managed? (n=764)		0.014	0.12	0.01
How has the school management changed compared to last year? (n=764)		0.048	-0.11	0.05
How would you rate the performance of school head teacher? (n=764)		0.064	0.16	0.05
Are you or a member of your family involved in CEC? (n=578)		0.28	1.3	0.6
Does the CEC have a regular communication with you about its plans and activities? (n=578)		-0.036	-0.17	-0.18
What kinds of actions or initiatives did this CEC take in the last 12 months? (n=578)	Monitor student attendance	0.18	0.27	0.55
	Monitor teacher attendance	0.01	-0.79	0.05
	Raise fund	0.41	-0.15	0.25
	Improve school infrastructure	0.49	0.56	0.53
	Support students financially	1.58*	1.6	1.59
	Buy learning materials	0.24	-0.35	0.54
	Promote enrolment of out-of-school children	0.06	-1.5	0.13

	Provide remedial support	0.21	-	0.11
	Reinforce the use of non-violent disciplines	0.14	-	0.37
	Monitor student retention	-0.62	-	-0.57
Did any of CEC's initiatives improve the quality of schooling your girls received? (n=578)		0.095	-0.03	0.13

The head teachers also received high ratings in the majority of cases (94.1 percent) for performing “excellent” (76.7percent) or “fair” (17.4 percent), although the SOMGEP-T program does not seem to have significantly improved any of these ratings across subgroups since the baseline (noting, however, the limited exposure time to the intervention – four months).

Although CECs exist in the majority of schools, only a small portion of the caregivers (17 percent) seem to interact with CECs as a member. Additionally, CECs appear to be involved and communicate with the parents about school-related issues in different frequencies, with 34.7 percent communicating on a monthly basis, 29.0 percent on a weekly basis and 3.4 percent on an annual basis, while 20.7 percent never inform the parents about plans and activities. In terms of the effect of the SOMGEP-T program, no significant increase was observed in the frequency of CECs’ communication with parents in the intervention group compared to the comparison group since the baseline.

When it comes to the actions and types of activities that CECs took in the last 12 months, monitoring student attendance happens to be the most frequent action cited by the parents (64.3 percent), followed by monitoring teacher attendance (51.0 percent), and improving school infrastructure (22.6 percent). However, provision of financial support to students is the only significant difference-in-difference among the intervention and the comparison groups. Put differently, the students from the intervention group appear to have received more financial support from their school’s CEC than their counterparts in the comparison group since the baseline. As the table below indicates, CECs in the intervention group have increased their financial support to students from 9.5 percent in the baseline to 16 percent in the midline, while it has dropped from 10.4 percent to 4.1 percent among the schools in the comparison group. No significant difference in differences were observed across subgroups in terms of the actions that CECs took in the last 12 months.

*Table 61: Provision of Financial Support to students in Midline- Household Survey*

	Baseline Intervention	Midline Intervention	Difference Baseline to Midline (Intervention)	Baseline Comparison	Midline Comparison	Difference Baseline to Midline (Comparison)	Difference-in-Difference
Provision of financial support (%)	9.49	16.02	6.53	10.41	4.15	-6.26	12.79

From the primary caregivers’ perceptions, the project does not seem to have reached its goal of increasing CECs’ monitoring of student attendance and student retention by 70 percent. Overall,

monitoring student attendance has increased from 52.5 percent in the baseline to 64.3 percent in the midline. The intervention group has had a 28 percent increase from the baseline (49.6 percent) to midline (63.5 percent) while the comparison group's percentage increase (15.9 percent) has been slightly less during the same period (56.5 percent and 65.5 percent, respectively). Unlike monitoring student attendance, the percentage of CECs monitoring student retention has decreased from 12.2 to 9.1 since the baseline. When data is disaggregated by intervention, most of this percentage decrease has been driven by the intervention group, in which only 8.31 percent of parents reported CECs monitoring retention in the midline compared to 13.9 percent in the baseline, whereas the CECs in the counterpart group's monitoring student retention has increased from 9.95 percent to 10.37 percent.

The findings above are consistent with the qualitative evidence gathered during FGDs. When asked about what mothers think about the role of CECs in changing girls' access to education, most of the mothers confirmed that CECs raise community awareness and support the girls from low-income families by paying their school fees. The majority of mothers seem to be well aware of the challenges facing CECs, such as financial constraints. For example, one of the mothers stated that "the CEC's main challenge is that there is no incentive for teachers to help students [...] and if the CEC members do not receive financial assistance, they really cannot do much else."<sup>240</sup>

However, there were mixed feelings among mothers about the role of CECs in improving school attendance. For example, some mothers said that CECs do not take any actions regarding students' absences while some others believed that students' attendance has generally improved due to the follow-up of CECs: "In the morning during school, the committee oversees students' attendance. Then they figure out those students who are absent and follow up by visiting their houses. The teacher's duty is to attend classes and provide lessons to students, not to oversee students' attendance. It is the CEC's job and they are responsible for figuring out missing students and sick students."<sup>241</sup>

Overall, more than half of the parents (51.3 percent) believe that CEC initiatives have improved the quality of schooling that their girls received. While insignificant, the positive coefficient of the difference-in-differences regression indicates that the parents of girls in the intervention group are more likely to believe that CEC activities are making any impact on the quality of their girls' education compared to their counterparts in the intervention group.

To summarize the findings, we analysed the role of CECs from the head teachers' and primary caregivers' perspectives – the head teachers in the intervention group provided a more negative assessment about the CECs' performance and the parents in the intervention group had a more positive assessment. Excluding the schools in the intervention group that receive significantly more financial assistance from CECs, none of these differences between the intervention and comparison groups were found to vary significantly across the school survey, household survey, and subgroups. The qualitative data revealed that CEC members may not have the adequate amount of skills to perform activities and participate in school management effectively which may be the reason why only half of the parents believe that the CEC is improving the quality of their girls' education.

Although the impact of the project was not evident across most of the school-governance related indicators due to the limited exposure time to the intervention, the project seems to meet the minimum standards of gender equality and social inclusion (GESI) framework. The level of data disaggregation used in this section is an indication of the project being GESI transformative and the management of the

---

<sup>240</sup> FGD- Mothers

<sup>241</sup> FGD- Mothers

SOMGEP-T project is interested to understand how the social, political and environmental factors affect girls' education and subsequently, incorporate those elements in the project design rather than focusing merely on girls' qualities.

### 6.3. Quality of teaching

One of SOMGEP-T's core interventions is a teacher training program that is designed to improve pedagogical techniques and overall teaching quality. Key activities include trainings for teachers on 1) improved delivery of literacy and English language, supported by digital content in all 148 primary and 55 secondary schools, 2) improved delivery of numeracy in all 148 primary and 55 secondary schools, and 3) to provide structured remedial support to students at primary and secondary level. A significant focus of the project is 1) the use of formative assessments – a type of continual evaluation that utilizes both formal and informal assessments of student ability and learning gaps – and 2) the reduction of corporal punishment in schools. Improvements in these areas are expected to have positive knock-on effects on learning and continued enrolment of girls in older grades.

However, because teaching quality is a broad, multidimensional concept, this section addresses several metrics of quality. Rather than define teaching quality by its downstream impact on learning outcomes, this section focuses on teaching quality as defined by teacher effort, teaching practices, and student and caregiver perceptions of teaching quality. The indicators of quality fall into five broad categories:

1. Caregiver perceptions of overall teaching quality, especially changes over the previous 12 months
2. The use of learner-centred pedagogy in class and the incorporation of formative assessments in teachers' processes
3. Classroom demeanour and safety: the use of corporal punishment in class, the extent to which teachers are respectful and welcoming toward students, and whether students feel safe and comfortable in the classroom
4. Gender equity in classroom interactions: Teacher's interaction in class with girls and boys e.g., the proportion of questions directed at girls, the use of harsh language directed at girls versus boys, and so forth
5. Teachers' effort level and preparedness for class, as indicated through the use of lesson plans and clear communication of learning goals at the start of the lesson

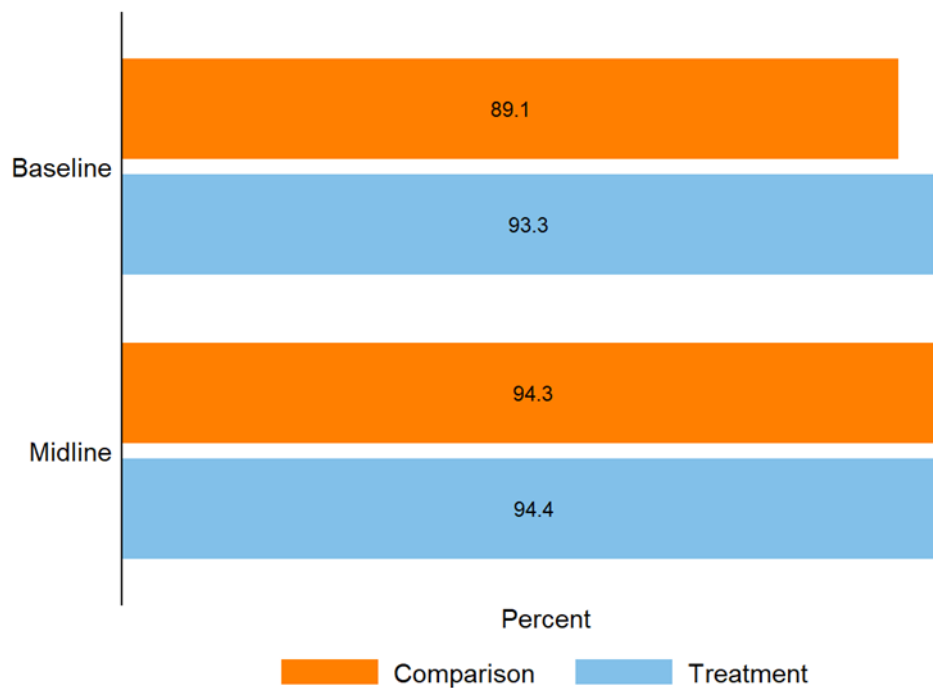
These indicators are captured using a combination of three distinct data collection tools. The primary tool is direct classroom observation by a Forcier researcher. Classroom observations were conducted in two classrooms per school at the baseline and one classroom per school at the midline. A researcher observed the classroom for a period of approximately 45 minutes, in three 15-minute blocks, recording details of the class, the level of participation, the use of corporal punishment, and other observations. The second tool is a survey of girls, as part of the household survey, which allows us to capture student perceptions of teaching quality, especially how comfortable students feel in class, and the use of corporal punishment when teachers are not being observed directly. Finally, the third tool was a survey of caregivers, implemented as a module of the household survey. Caregivers were asked about teaching quality in their girl's school over the past year.

#### Caregiver perceptions of teaching quality

Our first metric of teaching quality is also our broadest, and provides a useful summary of changes in teaching quality since the baseline. Caregivers were asked to assess the quality of teaching in their girl's school, and to indicate whether it has improved, stayed the same, or gotten worse over that period. As the figure below shows, caregiver perceptions of teaching quality have improved since the baseline – the

share of respondents who described their girl’s teacher as good or very good increased from 93.3 percent in intervention schools at the baseline to 94.4 percent in intervention schools at the midline. At the same time, caregivers also perceived an increase in teaching quality in comparison schools, and this shift – from 89.1 percent of caregivers describing teaching quality as good to 94.3 percent – was far larger than the shift observed in intervention schools. In terms of project impact, the improvement in comparison schools outstripped the improvement in intervention schools by 4.1 percentage points.

Figure 24: Share of caregivers who report good or very good quality teaching over past 12 months



It is possible that comparison schools, starting from a lower baseline, were simply catching up to intervention schools and that intervention schools – given their higher starting point – had little room for improvement. Unfortunately, this possibility is contradicted by two additional findings. First, when we redefine the question above and analyse the share of caregivers who perceive their girls’ teacher to be very good (as opposed to either good or very good), we still observe a larger positive shift in comparison than in intervention schools. Second, when caregivers were asked to assess the *change* in teaching quality in their girls’ schools over the past 12 months, comparison and intervention schools fared similarly: 85.7 percent of respondents in intervention schools reported that teaching quality had improved over the past 12 months, compared to 85.2 percent of respondents in comparison schools. While this result is subtly different from the figure above – insofar as the previous results suggest that comparison schools improved *more* than intervention schools and the current results suggest that they both improved equally – the broader point that intervention schools did not improve at a faster pace than comparison schools remains.

The following sections, which assess teaching quality on a wider variety of measures, will generally accord with the perceptions of caregivers reported here – intervention schools improved from baseline to midline, but not as much as comparison schools. Given that the program has been underway for less

than a year, and the sample size for many of our measures is small, it is not altogether surprising that many of the results are noisy and do not show significant project impact.<sup>242</sup>

The qualitative results suggest that teaching quality has improved, but that significant challenges still remain. Many qualitative interview participants discussed the positive changes to teaching quality that have resulted from the project, particularly the teacher trainings and capacity building activities conducted with local education stakeholders. As one teacher reports, “There is big change between last year and this year in the [...] school. Last year, students and teachers were not interacting with each other, but this year, students and teachers interact with each other and this caused teachers to attract students, and the teaching quality has improved.”<sup>243</sup> However, teachers discussed a number of challenges they face that are linked to teaching quality, including unfavourable teacher-student ratios, lack of training, lack of defined curricula, lack of materials and infrastructure, and low salaries/lack of salaries. These factors will also be discussed in the sections below to provide a full picture of which challenges the project has addressed, which challenges remain, and how the project can best adapt to remaining challenges.

### Learner-centred pedagogy

When teachers implement learner-centred pedagogy, students take a more active role in their own learning and teachers recognize that learners have different needs and abilities. We use two overall metrics of learner-centred pedagogy: the extent to which classrooms are active and participatory, and the use of formative assessments.<sup>244</sup>

Active and participatory classrooms come in many forms; in active and participatory classrooms, students engage in learning activities and learning material. They participate in classroom discussions, ask questions and are not passive learners who imbibe and memorize information shared by the teacher. We break this broad concept into eight distinct indicators, which reflect whether a classroom is participatory, e.g. if learning activities are being used, if students are observed instructing each other, etc. Trained team leaders recorded a simple “yes” or “no” for each indicator, capturing whether a particular participatory activity or outcome was observed during the observation period. A list of these indicators, as well as baseline and midline results for each, is provided in the table below.

As the table shows, there was significant improvement on these indicators in both intervention and comparison schools. For instance, in intervention schools, the share of teachers who asked students open-ended questions and asked for student opinions increased from 61.5 and 60.0 percent at baseline to 78.1 and 75.0 percent at midline, respectively.

*Table 62: Classroom participation and learner-centred methods, by intervention status and round*

	Baseline		Midline		DID
	Comparison	Intervention	Comparison	Intervention	

<sup>242</sup> Sample size for class observations in midline was one observation per classroom block, resulting in 62 observations in total. At the baseline, two observations were completed per school, resulting in a baseline sample size of 124 observations (186 in the combined baseline and midline sample).

<sup>243</sup> FGDs – Teachers

<sup>244</sup> Formative assessments are assessments where teachers use results from assessments to inform their lesson plans – both what is taught and how it is being taught. Formative assessments include written and oral tests as well as learning activities.



Did NOT spend most time copying from board	72.9%	86.2%	93.3%	93.8%	-12.8%
Did NOT spend most time repeating words aloud	83.1%	84.6%	100.0%	100.0%	-1.5%
Use of student-centred games/activities	39.0%	30.8%	40.0%	40.6%	8.8%
Students instructing each other	37.3%	41.5%	40.0%	56.2%	12.0%
Teacher asks open-ended questions	44.1%	61.5%	70.0%	78.1%	-9.3%
Teacher asks for student opinions	57.6%	60.0%	80.0%	75.0%	-7.4%
Sought to involve student who was not participating	69.5%	60.0%	66.7%	81.2%	24.0%
Students worked in groups	42.4%	32.3%	23.3%	25%	11.8%

\*\*None of the DID (difference-in-differences) estimates of project impact are statistically significant

The findings in the table are extremely consistent with respect to improvements in intervention schools: intervention schools improved on every metric reported here, except the use of group work, and typical improvements were sizable. These improvements in intervention schools were smaller than those observed in comparison schools in the case of reducing passive teaching methods – such as not copying from the board, not repeating the teacher’s words aloud, asking students open-ended questions, and asking for student opinions.<sup>245</sup> However, the greater relative gains in comparison schools is in part due to the lower base from which comparison schools started in the baseline. The percent of teachers who did not have students spending most of their time copying from the board, did not have student repeating words aloud, and who asked open-ended questions was lower in comparison schools than in intervention schools. While comparison schools did improve in the midline along these measures, they did not surpass or only matched intervention schools in absolute terms. Intervention schools improved absolutely and also relative to comparison schools in the use of student-centred activities, in the extent to which teachers allowed students to instruct one another, and in seeking to involve students who were not participating in the class.

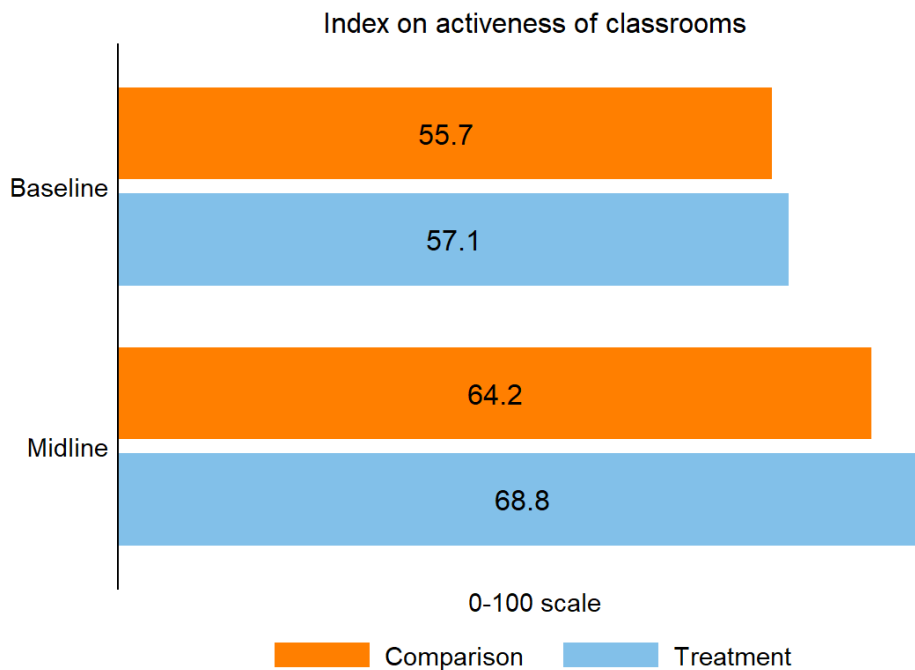
Because no single measure captures the full diversity of classroom participation, we constructed an index that combines the eight indicators into a single score, ranging from zero to one.<sup>246</sup> The performance of intervention and comparison schools, disaggregated by evaluation round, are shown in the figure below. Consistent with the results from the eight individual indicators reported previously, teachers in both

<sup>245</sup> As is the case in several other areas throughout this report, it is difficult to determine the source of improvements observed in comparison schools without significant additional data collection. In the context of teaching quality, comparison schools could be assisted by the actions of other NGO training programmes. But other explanations are also possible – improvements in the conflict environment assisting teacher retention, or even random sampling variation. In future evaluation rounds, we encourage the use of additional, targeted questions to head teachers and parents to determine what other events or changes of note to the evaluation have occurred in their communities.

<sup>246</sup> The index was created by taking the mean of the eight measures of classroom activity levels. Each of the eight measures were measured on a 0/1 binary scale, with “1” being more active. The first two measures, copying from the board and repeating words aloud were reversed to negative statements (e.g., did NOT copy from board), so that 1 indicated greater activity levels for all eight measures.

intervention and comparison schools showed improvement in this aggregate index. Intervention schools improved more than comparison schools in absolute terms, but this gap was not statistically significant when assessed in a formal difference-in-differences framework.

Figure 25: Index of active and participatory classrooms, by intervention status and round



Beyond the use of participatory methods in class, a core project goal is to increase the use of formative assessments. Either before or immediately after the classroom observation began, teachers were asked whether they use formative assessments in their teaching. Those teachers who indicated that they do use formative assessments were asked whether they had records or documentation of their use.<sup>247</sup>

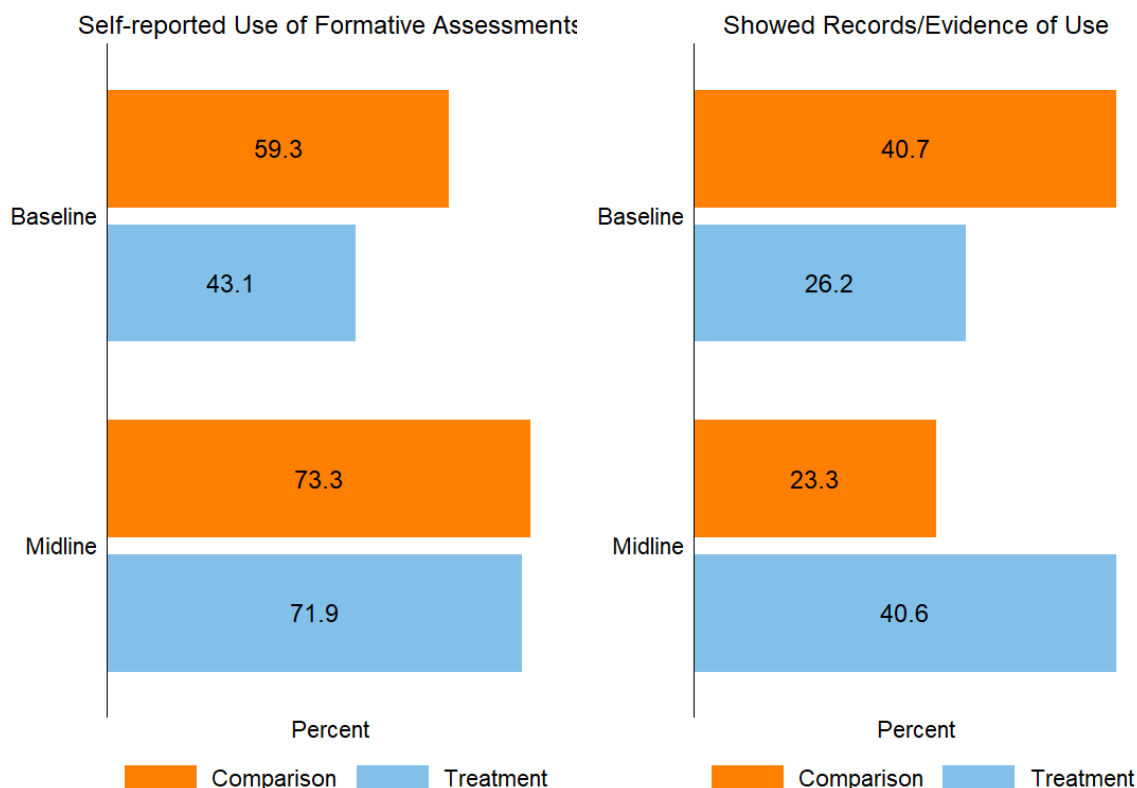
In the figure below, we report the share of teachers who claim to use formative assessments and the share of teachers who have documentation of their use. The results indicate that teachers are becoming more learner-centred in their classrooms.<sup>248</sup> Results show that the use of formative assessments increased across the board, but that gains in intervention schools were much larger overall – relative to comparison schools, teachers in intervention schools saw a 14.8 percentage point increase in the use of

<sup>247</sup> Teachers were not specifically asked to produce this documentation for two reasons: first, the classroom observations were being conducted during class time, and researchers did not want to disrupt the class while teachers sorted through their files. Second, because teachers were not told in advance that they would need to furnish proof of their use of formative assessments, most would not be prepared to do so. Instead, we simply accept, *prima facie*, teachers’ claims regarding their possession of records.

<sup>248</sup> Formative assessment refer to the use of ongoing assessments (using activities, oral examinations and written tests) by teachers to adapt learning to meet the need of learners. Formative assessment recognizes that learning is a journey and not a “pass/fail” state, and that each learner might have different strengths and challenges in learning, which is an essential aspect of learner-centred pedagogy. Hence, teachers who use formative assessments often check for if and how students are learning, and change their teaching plan if necessary.

formative assessments (left panel) and an even larger increase of 31.8 percentage points in the share of teachers who report having documentation of their use of formative assessments.

Figure 26: Teachers' use of formative assessments



### Progress to Teaching Quality Target

The measure of progress for the teaching quality intermediate outcome is improvement in the proportion of teachers who apply improved teaching practices in literacy and numeracy. This measure is formally defined as the proportion of teachers sampled who apply formative assessments among all teachers sampled. The project's target for the midline evaluation is that the percent of teachers who report using formative assessments will increase by 10 percentage points over the percent reported in the baseline. In the baseline evaluation, only 43.1 percent of teachers reported use of formative assessments in their classroom. As such, the target for the midline is 53.1 percent of teachers. In the midline evaluation, teachers in the intervention schools far surpassed that target. Among all teachers in intervention schools, in the midline, 71.9 percent of teachers reported that they used formative assessments, an improvement of 28.8 percentage points beyond the target.

The teaching quality target for the midline is also met if we consider the proportion of teachers who have records showing that they use formative assessments. In the baseline, 26.2 percent of all teachers in intervention areas had records of using formative assessment with their students. Therefore, the target for the midline for this measure of teaching quality is 36.2 percent of teachers. In the midline evaluation, we

find that 40.6 percent of teachers have evidence of formative assessment usage, 4.4 percentage points above the midline target.

There is evidence in the qualitative data that teachers have begun using formative assessments since the start of the project. Teachers appear to be using both formal and informal assessments to gauge students' progress and test their understanding of content. For example, when asked whether teaching methods have changed at all in the past year, one teacher responded, "Yes, we have changed. We now conduct monthly tests or tests at the end of a chapter."<sup>249</sup> Teachers use less formal methods as well, such as question and answers sessions, when relevant; there is an increased focus on tailoring methods used to different students' needs. One teacher described this process of using different methods to cater to students' needs: "All students are different, some understand better with pictures, some with questions and answers – therefore, the teacher needs to understand the needs of each student."<sup>250</sup>

In the project logframe, there was one qualitative indicator for teaching quality: Shifts in teachers' awareness of quality of education. The baseline target for this indicator was "Child protection remains a major issue; corporal punishment is widespread and associated with students who do not understand the content. Gendered practices observed in math teaching." The midline target was "Teachers are aware of the need to support students who are lagging behind in acquiring literacy and numeracy skills; identify sub-groups who are struggling and potential strategies for inclusion." KIs with teachers were the primary data source for the analysis of this indicator.

In the SOMGEP-T baseline, there was already some awareness among teachers on the negative impacts of corporal punishment, but little evidence of child protection policies or formal strategies for handling child protection cases. Girl and boy students in the baseline largely reported that corporal punishment is used on students who are particularly disruptive. However, there was also some evidence that teachers used corporal punishment on students who do not come to class prepared or have trouble understanding the content. Although there was no direct evidence of gendered practices in math teaching in the qualitative data, teachers across FGDs frequently described girls as shy and less active than boys in class.

The midline qualitative data suggests that understanding of the negative impacts of corporal punishment has increased among teachers, and teachers appear far less likely to use corporal punishment on students who do not understand the content of lessons. In most cases, teachers explicitly stated that use of corporal punishment on girls has a negative, counterproductive effect. As one teacher explains, "If you are a teacher, you cannot beat the student if the student does not understand something. It may have happened when we were students, but the situation is different now. Teachers are receiving lots of training about how students understand something."<sup>251</sup> Another explains that if you use corporal punishment "the student might run away from the school, but that is not good. It is good for the student to receive private one-to-one lessons."<sup>252</sup> In the few cases where teachers admitted to using corporal punishment, it appears as though corporal punishment is used as a last resort for students who are persistent in misbehaving. For example, "Some students do not take advice when you give it. Students

---

<sup>249</sup> FGD – Teachers

<sup>250</sup> FGD – Teachers

<sup>251</sup> FGD – Teachers

<sup>252</sup> FGD – Teachers

that do not behave well and do not take advice will be beaten because they should not come to school with their bad manners.”<sup>253</sup>

In terms of teaching practices, teachers appear to be making an effort to engage and make special allowances for disadvantaged sub-groups. Generally, there is an awareness of the unique demands girls face in their households, as well as the need for specialized teaching plans and treatment to help girls cope with these demands while still succeeding in school. When asked whether there is a difference between how boys and girls participate in the classroom, one teacher responded, “There is a difference. Boys are more active than the girls because girls are always doing chores and they do not have the chance to read books. To uplift the girls, I always motivate them and they are getting better compared to previous years.”<sup>254</sup> Another teacher explains that there is a need for private or repeated lessons with girls who struggle to participate: “Sometimes girls are shy so I repeat lessons with female students exclusively.” Teachers also make special arrangements for students who live far or face difficulties regularly attending or affording school. As one teacher explains, “Students are not the same and come from different distances. Some students need to travel up to three hours to get to school while others only have to travel less than an hour to get to school. We understand and accommodate for the individual circumstances of each student.”

Teachers also appear to be increasingly equipped to handle students who are in the same grade but at different levels of learning: “The students are not same; some of the students have open minds and they can understand quickly, some of them can understand the lesson but they have bad handwriting, some have good handwriting but their comprehension is very bad. So the teacher is the one who can motivate the student; last year I bought books to make lessons with drawing. I tried to give students a lot of motivation. I gave awards to the students who won first in the class; this motivated other students to make an effort.”<sup>255</sup> Another teacher describes having shifted to dictating lessons instead of writing them, as well as deploying special teaching methods to include both boys and girls in classroom activities: “I dictate the Somali lessons to the students instead of writing on the board. I always select five boys in one day while five girls are selected the next day.”<sup>256</sup> Other methods teachers mentioned using to include students who are struggling include holding competitions, moving students from the back of the classroom to the front and encouraging students who are in the back to participate, repeating lessons, following up with questions to ensure students understand lessons before moving on, and following up with parents to ensure they are supporting their child at home.

There does, however, appear to still be a serious gap in child protection. Neither teachers nor students discussed formal child protection strategies or policies, and there was extensive evidence from girls with disabilities and mothers suggesting that children with disabilities face so much physical and verbal abuse in school that they are discouraged from attending or are forced to take matters into their own hands. One girl with disabilities explained in her interview the abuse she faces in school and how she retaliates: “When we are in the school, some of the students call me ‘the girl who only has one eye.’ Then I attack them and they stop insulting me.”<sup>257</sup> When asked what the school does to respond to this abuse, she replied, “The school separates us, but they do nothing else.” It appears that very little has been done to address this abuse at the school level. Only a few teachers reported that their school helps disabled students, and actions appear limited from a child protection standpoint – providing them with materials,

---

<sup>253</sup> FGD – Teachers

<sup>254</sup> FGD – Teachers

<sup>255</sup> FGD – Teachers

<sup>256</sup> FGD – Teachers

<sup>257</sup> IDI – GwD

moving them closer to the board, discouraging other students from insulting them, and providing them with one-on-one learning sessions in the afternoon.

## Classroom Environment

For students to learn effectively and desire to stay in school, it is important that they feel comfortable at school and that the classroom serves as a positive learning environment. A safe and comfortable classroom environment encompasses several different aspects of teaching, including a teacher's demeanour and the disciplinary policies in their classroom. We assessed the nature of the classroom environment using several indicators, including: reported and observed use of corporal punishment, the manner in which teachers engage with their students (i.e. how respectful they are), and how welcome students feel in class. This is measured through a combination of classroom observations and student interviews.

Students who feel safe in their classroom are more likely to participate in discussions and ask questions. They are also more likely to enjoy attending school and hence student attendance – and potentially continued enrolment – is likely to be higher. This indicator is very important for girls, especially, who might feel less comfortable in classrooms and less sure of whether they belong, especially if the majority of the teachers and adults in the school are male.<sup>258</sup>

One finding that emerged from the qualitative data was that positive student-teacher relationships are particularly important in schools that have unfavourable teacher-student ratios. In the learning outcome section, we briefly discussed the disproportionately large effect a single teacher can have on girls' education outcomes in schools that have unfavourable teacher-student ratios – in the qualitative data, teachers attest to the fact that girls are shy in the classroom and hesitate to ask for help when they do not understand material, and multiple REOs mentioned that they are trying to recruit female teachers to improve girls' retention. Additionally, mothers and girls themselves provide direct evidence of the importance of positive student-teacher relationships. One mother explains the link to attendance: "Girls don't attend school when they are scared or when they have their period." In the qualitative risk mapping exercise, one girl indicated that she is scared of the teachers and that she does not feel safe in the school office. She explains her reasoning: "Because whoever did something wrong will be punished in there. And also as girls we feel uncomfortable being there. We feel shy."<sup>259</sup>

Another way in which unfavourable teacher-student ratios negatively affect teaching quality is through forcing principals to pivot away from management tasks (e.g. monitoring attendance, school plans, and personnel) and take on a teaching role. In some schools, a history of strong diaspora involvement appears to have left a legacy, as these schools have been able to retain good quality teachers by offering stable salaries. One teacher explains the direct effect diaspora support has on teacher retention: "Firstly, this school was established by diaspora from the outside and they have been supporting this school (teachers and management). Last year, they supported the activities to cope with the droughts in the community. The diaspora do not support anymore - now the number of the teachers has reduced."<sup>260</sup> However, in other schools, this support is lacking: "Not all teachers receive salary and because of that some leave the school."<sup>261</sup>

---

<sup>258</sup> The teacher corps is dominated by male teachers overall. Of 388 full-time teachers documented at the 63 schools visited during the midline, only 15.2 percent were female.

<sup>259</sup> Risk Mapping – Girls

<sup>260</sup> FGD – Teachers

<sup>261</sup> FGDs – Mothers

Findings regarding the use of corporal punishment were mixed, though most of the results suggest that corporal punishment declined in intervention schools from baseline to midline. For instance, Forcier researchers observed the use of corporal punishment in 66.2 percent of intervention classrooms at the baseline, but just 31.3 percent at the midline. While this measure faces methodological concerns stemming from the effect of a classroom observer and social desirability bias – i.e. the Hawthorne effect, in which participants in a research study behave differently, because they know they are being observed, compounded by the fact that many teachers are aware that corporal punishment is considered unacceptable – the social desirability bias was present in both baseline and midline evaluations and declining prevalence of corporal punishment is still noteworthy.<sup>262</sup>

In addition, the prevalence of corporal punishment declined in other measures as well. As shown in the table below, the share of students who report that their teacher used physical punishment *on any student* in the last two weeks declined by 13.2 percentage points in intervention schools from baseline to midline. And, although the broader question of physical punishment (which asked students whether their teachers use physical punishment at all, as opposed to whether they had used it in the last two weeks) shows an increase at midline, the context in which this question was posed also changed between the two evaluation waves.<sup>263</sup> As a result, we believe it is preferable to focus on the results derived from direct observation and from student recall over a specific time period, both of which registered declines in the use of corporal punishment in intervention schools. Asking girls whether their teachers used corporal punishment *within the last week* avoids the social desirability bias involved with direct observation of teachers; avoids relying on a measure – like general reports, by girls, of corporal punishment use in their classrooms – whose construction and interpretation by respondents has changed substantively since the baseline; and provides a specific recall period over which girls are asked about corporal punishment, which is preferable from a question design standpoint. That this measure accords, broadly, with the findings from direct observation as well provides additional confidence in the result.

*Table 63: Use of corporal punishment, by intervention status and round*

	Baseline		Midline		DID
	Comparison	Intervention	Comparison	Intervention	

<sup>262</sup> An additional methodological concern stems from the subjective definition of corporal punishment. Forcier team leaders were trained to identify corporal punishment and defining when discipline crosses the line into physical punishment was discussed explicitly during training. More importantly, however, the same set of team leaders conducted the classroom observations at baseline and midline. Therefore, if their personal, subjective approaches to identifying physical punishment during classroom observation is stable over time, there is no reason to expect that enumerator bias would explain the large decline in corporal punishment rates over time.

<sup>263</sup> At the baseline, students were asked whether their teacher punishes students who “get things wrong in a lesson.” As a follow-up question, they were asked what form of punishment their teacher uses. The context of the question, therefore, implied that they should describe the form of punishment teachers used when students make a mistake in their lesson – one specific disciplinary scenario. Additionally, only some students were asked the follow-up question, if they indicated that their teachers punish students who make a mistake in class. In contrast, the midline reordered questions so that students were asked how teachers discipline students without reference to a specific context or type of trouble, and all students were asked this question. The differences in both the subsample who responded to the question and the contextual backdrop suggest that comparisons between baseline and midline on this question should be considered tenuous at best.



Observed use of physical punishment (toward any gender) <sup>264</sup>	89.8%	66.2%	20.0%	31.3%	35%
Observed use of physical punishment (toward girls) <sup>265</sup>	86.4%	64.6%	16.7%	28.1%	33%
Students report - use of corporal punishment by their teachers in class	52.7%	43.0%	62.9%	55.8%	3%
Student report - Teachers discipline or punish students who get things wrong in a lesson	76.5%	76.0%	86.3%	80.8%	-5%
Student report - Teacher used physical punishment on other students in last week	50.9%	64.9%	53.2%	51.7%	-16%

\*\*None of the DID (difference-in-differences) estimates of project impact are statistically significant

While the use of corporal punishment appears to have declined in intervention schools over time, we observe a very large decline in the use of corporal punishment in comparison schools as well. In the case of direct observation, corporal punishment was observed in 89.8 percent of classrooms in comparison schools at the baseline and just 20.0 percent at the midline, implying that intervention schools – though registering a decline themselves – actually lag behind comparison schools on this metric. At the same time, comparison schools do not show a decline in student reports of corporal punishment used in the previous week.

In the qualitative interviews, most teachers reported changes to punishment methods – away from physical punishment and toward use of warnings and private counselling, particularly for girls. Physical punishment was only explicitly mentioned in a few interviews, and in many other interviews teachers explicitly stated that use of corporal punishment can have severe negative effects on students. One teacher explains the importance of creating a safe classroom environment: “A teacher is like a second father to the student. Therefore, I show them that I'm like their parent. I talk to them, I welcome them, and I make them feel confident if they ask me a question. But if you are very strict, they will be too afraid to ask, even if the lesson is difficult for them. There is a big difference between when I first came to the school and now. Right now, the students and I understand each other well.”<sup>266</sup> Another teacher explains, “It is better to help a student understand by talking to him/her. Beating a student does not produce a well-disciplined student.”<sup>267</sup> Teachers also reported communicating with parents and school administration in order to find solutions when children misbehave.

*Classroom demeanour of teachers*

The classroom environment encompasses more than just the use of corporal punishment. Teachers can make students feel welcome and comfortable in class beyond simply refraining from the use of physical punishment. For instance, teachers should address students by their name, use respectful language, ensure that all students are given a chance to participate, and provide encouraging feedback when they do. To assess the classroom environment, researchers observing their classrooms recorded whether

<sup>264</sup> If girls or boys were observed to be disciplined physically in class.

<sup>265</sup> If girls were observed to be disciplined physically in class.

<sup>266</sup> FGD – Teachers

<sup>267</sup> FGD – Teachers

teachers used respectful language toward students, appeared to know and use students' names consistently, and avoided the use of harsh language when speaking with students.

Intervention schools improved in measures of classroom demeanour approximately as much as comparison schools. As shown in the table below, 84.6 percent of teachers in intervention schools in the midline did not use harsh language when interacting with students, compared to 81.3 percent at the baseline.<sup>268</sup> This 3.3 percentage point improvement is comparable with the 3.1 percentage point improvement in comparison schools.

Our other indicator of change over time focuses on students themselves and whether they report feeling welcome in the classroom. This is a particularly useful metric, because it captures all of the factors that go into making a student feel welcome or unwelcome. As with the previous results, both intervention and comparison schools improved over time – at baseline, 77.1 percent of students agreed strongly with the statement “my teachers make me feel welcome in the classroom,” and this share increased to 84.7 percent at the midline. These gains were especially concentrated among intervention schools, which improved by 11.1 percentage points, while comparison schools improved by just 3.4 percentage points.

*Table 64: Teachers' classroom demeanour, by intervention status and round*

	Baseline		Midline		DID
	Comparison	Intervention	Comparison	Intervention	
Harsh language <b>not</b> observed	83.3%	81.3%	86.4%	84.6%	0.2%
Students report feeling welcome by teachers <sup>269</sup>	77.1%	77.2%	80.5%	88.3%	7.7%**
Respectful language observed <sup>270</sup>			76.7%	65.6%	NA
Teacher used student names			46.7%	56.3%	NA

\*\*Difference-in-differences estimate of project impact is significant at 10% level.

Our remaining measures of classroom demeanour and environment were introduced in the midline evaluation wave, which precludes over-time comparisons. However, the results, shown in the bottom two rows of the table above, show mixed results between intervention and comparison schools. Teachers in midline schools were largely reported to use respectful language for students (71.0 percent were very or somewhat respectful) though only about half (51.6 percent) knew their students' names.

<sup>268</sup> In class observations, harsh observations were reported on a Likert scale of: None observed 1 time observed 2+ times observed. This was converted to a binary of observed/not observed

<sup>269</sup> Students reported if teachers made them feel welcome in the classroom. All negative and missing values taken as “not welcoming.”

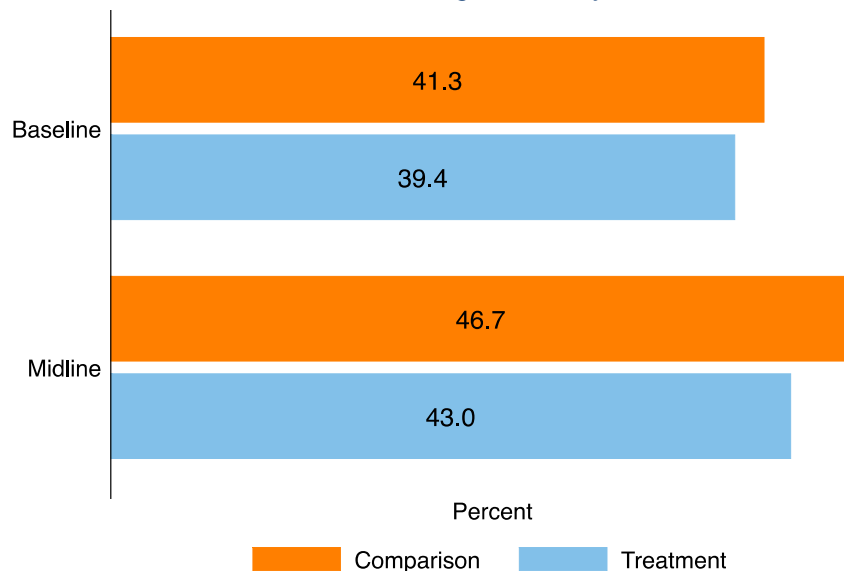
<sup>270</sup> In class observations, teacher's tone was recorded on a Likert scale of: Not respectful, Somewhat respectful, Very respectful, Respectful. This was converted to a binary scale where not respectful and somewhat respectful are taken as “not respectful.”

## Gender equity in classroom interactions

Gender equity in teacher interactions with girls and boys in the classroom is an important factor for increasing girls' participation in classes and preventing them from dropping out. Gender equity in access to educational opportunities alone is not sufficient – bias within classrooms, either explicit or implicit, can discourage girls from staying in school and limit their learning opportunities. A conducive learning environment for girls includes equity in their opportunities to participate in class, interact with their teachers, and receive feedback on their work. This section reports indicators on gender equity in classroom participation as well as in how teachers interact with girls and boys in the classroom.

As an initial test of gender equity and changing levels of equity over time, students were asked whether they feel boys and girls are treated differently by their teachers. The share of students (all female) who report that boys and girls are treated differently increased from baseline to midline, in both intervention and comparison schools. This increase was 2 percent higher in intervention schools. However, it is important to emphasize that this increase could represent either a positive or a negative development with regard to girls' opportunities in the classroom, due to the ambiguous nature of the question: girls could view themselves as being discriminated against, and indicate that girls and boys are treated differently; at the same time, girls could be given special attention in class or additional opportunities to participate, and they might, therefore, report that they are treated differently (but in a positive manner) than boys.

Figure 27: Share of female students who believe girls and boys are treated differently at school



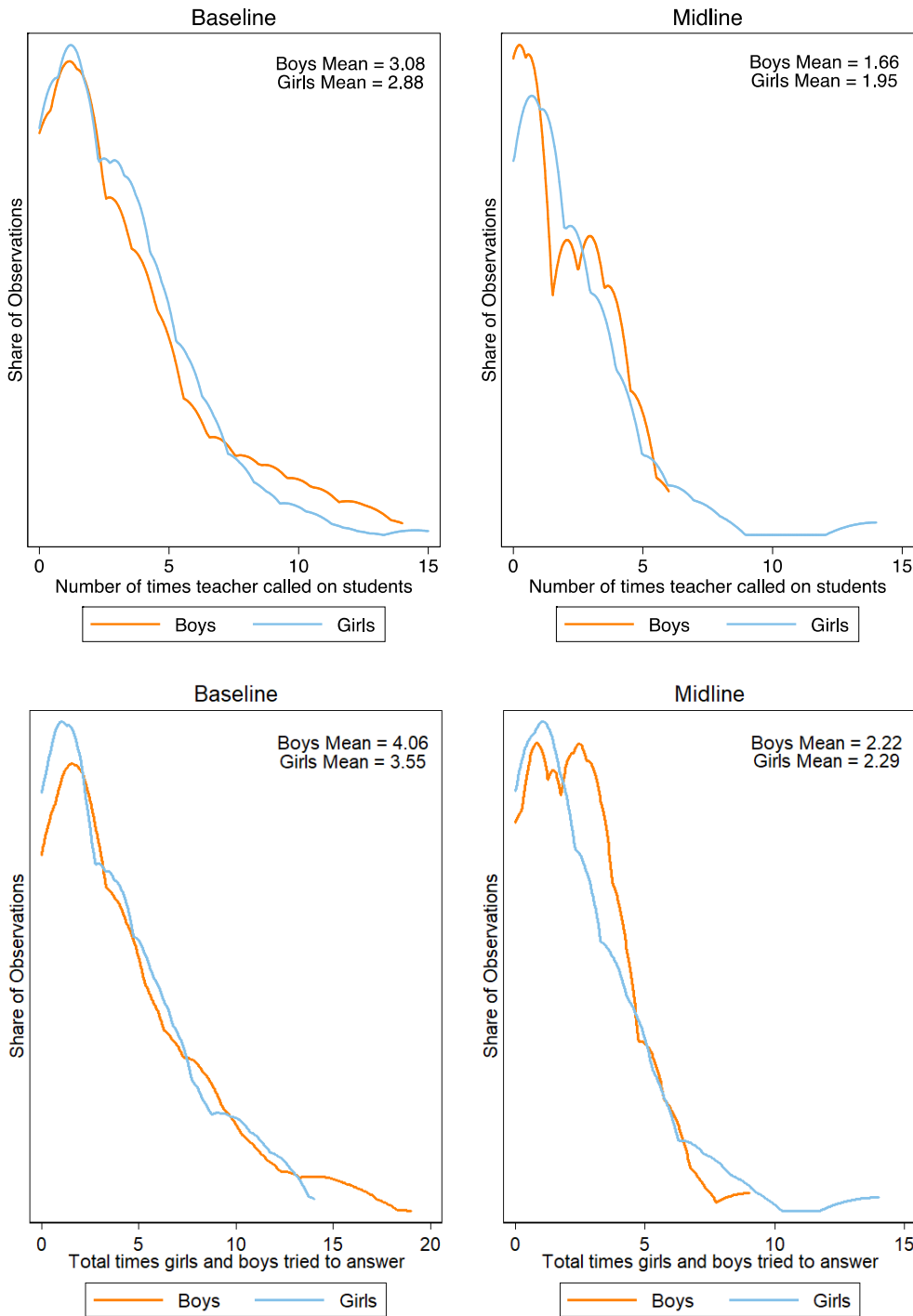
To further investigate gender equity in the typical classroom, we reviewed four additional indicators, all derived from direct observation of teachers in their classrooms:

1. The number of times female students tried to answer a question in class, versus the number of times male students tried to answer questions in class
2. The number of times the teacher calls on female students, compared to the number of times they call on male students
3. Number of times that the teacher provides encouraging feedback to female versus male students
4. Number of times that the teacher uses a harsh tone toward female versus male students

During the classroom observation, enumerators recorded the number of times the teacher called on male and female students during 15-minute blocks. In the graph below (top panel), we plot how many times teachers called on male and female students during their blocks at the baseline and midline. At the midline, the typical teacher called on male students 1.66 times during a 15-minute block and called on female students 1.95 times during the same period. The number of times a teacher calls on students declined overall from baseline to midline. However, this need not be concerning as this could be because teacher has shifted to using learning activities and learner-centred pedagogy, where teachers would ask fewer questions and would encourage them to participate themselves.

While the results in the top panel of the figure below show an increase in the frequency with which girls are called on relative to boys, the change from baseline to midline was not statistically significant. Notably, the shift toward girls' participation from baseline to midline was strongest in intervention schools, but this difference is not sufficiently large, to distinguish from a null effect. The results in the bottom panel – which focus on the number of times girls and boys, respectively, attempted to answer a question in class – show a similar pattern: an aggregate decrease in the number of attempts from baseline to midline, and a moderate shift toward greater girls' participation.

Figure 28: Frequency of classroom participation by boys and girls, by round



The evaluation also included questions on how often male and female students receive positive, encouraging feedback from their teachers. At the baseline, encouraging feedback to a boy was

documented in 56.4 percent of classrooms, and encouraging feedback to a girl was documented in a similar share, 57.3 percent, of classrooms. These rates rose at the midline, to 61.3 and 62.9 percent, respectively, as shown in the table below. Notably, the single largest shift in teacher encouragement was in intervention schools and specifically with regard to girls: encouragement of girls rose 11 percentage points from baseline to midline in intervention schools, compared to a 1-point increase in comparison schools, consistent with a more positive environment for girls' participation in intervention schools over time.<sup>271</sup>

*Table 65: Teachers' use of encouraging feedback for boys and girls, by intervention status and round*

	Baseline		Midline		DID
	Comparison	Intervention	Comparison	Intervention	
Teacher provides encouraging feedback to <b>boys</b>	61%	52%	67%	56%	-2%
Teacher provides encouraging feedback to <b>girls</b>	59%	55%	60%	66%	10%

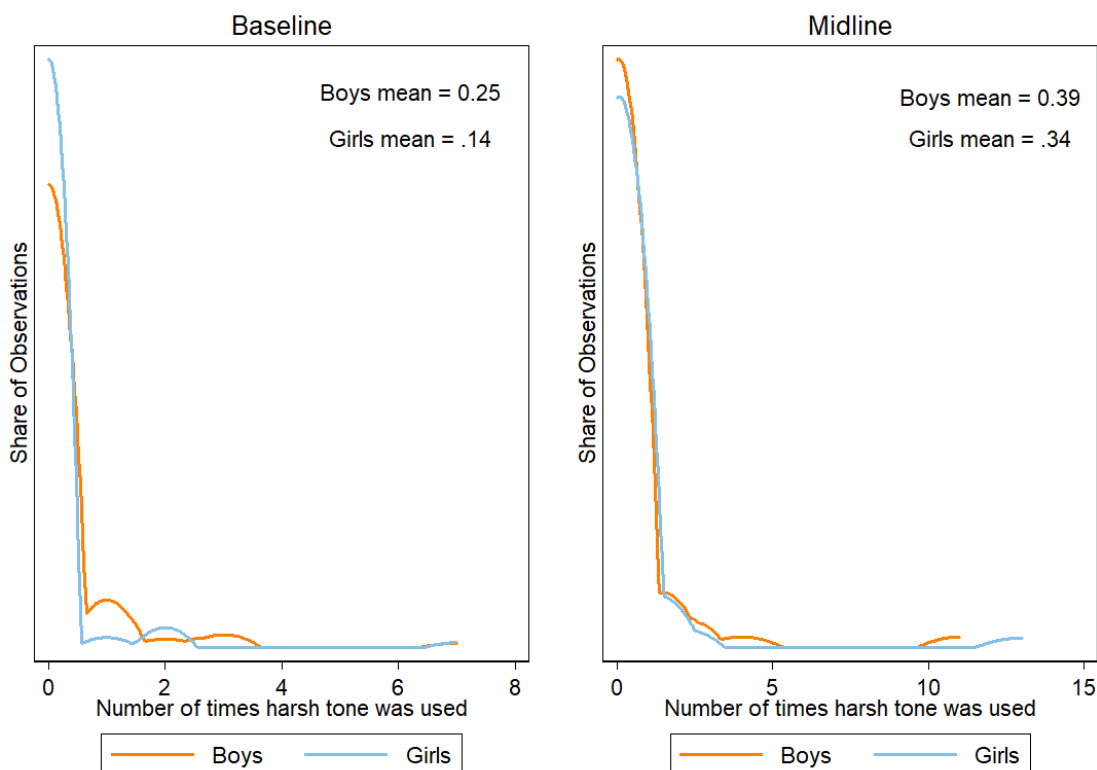
\*\*None of the DID (difference-in-differences) estimates of project impact are statistically significant

While the use of positive, encouraging feedback increased overall from baseline to midline, across the full sample of schools, so did the mean number of times that teachers used a harsh tone in class. When directed at boys, the mean count increased from 0.25 times per observation period to 0.39 times, as shown in the figure below. An even larger increase occurred when considering teachers' use of a harsh tone toward female students, with female students nearly reaching par with male students by the time of the midline. Worryingly, the net increase in the use of harsh tone and, especially, the increase in the use a harsh tone directed at girls, were concentrated among intervention schools. In these schools, the mean count of harsh comments directed at girls increased from 0.09 per period to 0.53, surpassing the increase seen with regard to boys.

It is possible that increasingly harsh verbal interactions with students are an outgrowth of shifts away from the use of corporal punishment. Teachers need to maintain order in their classrooms; in the absence of effective alternatives to corporal punishment, they may feel that harsh verbal punishment is their only recourse, particularly if they have not been adequately trained or have not internalized alternative approaches to classroom discipline. Unfortunately, the data available from girls is only suggestive: girls report an increase from baseline to midline in the share of teachers who shout in class as a form of discipline or a method to maintain order; however, this measure does not incorporate a specific recall period and the context in which it was asked has changed from baseline to midline, so we caution against relying too heavily on the results.

<sup>271</sup> Note, however, that this shift was not statistically significant at any conventional level in a difference-in-differences test.

Figure 29: Teachers' use of a harsh tone with boys and girls, by round



When teachers and parents were asked in the qualitative interviews whether there are differences in the participation and treatment of boys and girls in the classroom, they gave varied responses. As in the baseline, the most common response was that girls are shy in the classroom and that, as a result, they do not actively participate in class or alert the teacher when they fail to understand lessons. As one teacher explains, “In my opinion, there is a difference in classroom participation between boys and girls. Girls know all of the answers, but they are shy to speak out compared to boys. Even when boys know an answer is wrong, they will still participate in class.”<sup>272</sup> Additionally, some teachers and parents explained that girls may have more trouble participating in class because they are burdened by household chores, which prevents them from doing their homework or attending class regularly. However, there were a number of respondents who believe girls are more ambitious, focused, and active in school activities than boys. One teacher has observed significant changes in the classroom: “Most of the time girls are shy, but that does not mean that there is a lack of knowledge. There was a time when boys were more active than the girls, but right now everything has changed. We can now have competitions between boys and girls [thanks to] social media, mobile phones, and TV. These competitions have become helpful.”<sup>273</sup>

Many teachers appear to understand the importance of encouraging girls' participation. One teacher explains, “Boys are more active than the girls because girls are always doing chores and they do not have the chance to read a book. To uplift the girls, I always motivate them, and they are getting better

<sup>272</sup> FGD – Teachers

<sup>273</sup> FGD – Teachers



compared to previous years.”<sup>274</sup> Another explains that he encourages girls to participate and encourages other teachers to do the same and give them a chance to participate more in class. However, not all teachers understand how to encourage girls or appear to be motivated to do so. When asked whether there have been any changes with girls’ participation in the classroom, one teacher replies, “Girls are always like this. There is no interaction, and this is just the culture.”<sup>275</sup> These results suggest that there is still ample room for improvement in teachers’ treatment of girls.

### Teacher effort

Teacher effort refers to the energy, time and resources spent by the teacher to prepare and deliver lessons. Higher teacher effort is expected to translate into more structured lessons and higher teaching quality overall, and eventually into better learning outcomes. Our measures of teacher preparedness and effort were only reported in the midline – hence, we cannot say whether there have been improvements over time, nor whether the changes in intervention schools have been larger than those in comparison schools.

Our first measure considers the extent to which teachers are prepared for class and appear to have a lesson plan ready when they arrive. Lesson plans are an important tool for teachers to set learning goals, decide teaching methods and set a standard pattern of teaching in the course of the year. Overall, teachers in comparison schools were more prepared, according to the researchers conducting observations of their classrooms, than their counterparts in intervention schools – 57 percent of teachers in comparison schools were rated “very prepared,” compared to just 47 percent of teachers observed in intervention schools. Even more worrying is the share of teachers who were reportedly “not prepared at all”: only 3.3 percent of comparison school teachers fit this description, but 15.6 percent of intervention school teachers were described this way.

*Table 66: Teacher preparedness and effort level*

Indicator	Comparison	Intervention
Teacher seemed very prepared and had a plan for the observed lesson	57%	47%
Teacher clearly communicated the objective of the lesson	93%	78%

Our second measure of teacher preparedness and effort, reported in the table above, focuses on their communication of lesson objectives. Clear communication of lesson objective indicates that the teacher had a clear lesson plan. It also makes learning easier by helping students connect to what they are learning, orient what they are learning relative to past lessons, and keep track of their learning journey. By this metric, as with more general preparation levels, intervention schools lagged behind their comparison group: 93 percent of teachers in comparison were observed clearly communicating the lesson plans compared to 78 percent of teachers in intervention schools.

The qualitative data suggests that the combined effect of salary concerns, poor management, lack of adequate training, and lack of proper materials can lead to low levels of confidence and motivation among teachers. The teachers interviewed through the qualitative portion of the research most often cited mathematics and language courses (Somali and English) as being the most difficult courses for students.

<sup>274</sup> FGD – Teachers

<sup>275</sup> FGD – Teachers

One teacher explains, “Yes, there are some areas that are particularly difficult, and students struggle to understand them well, for instance Grade 7 and 8 Math, especially the chapters on algebra and geometry. When we teach geometry, we need additional tools that we do not have, and it becomes difficult for the students to understand.”<sup>276</sup>

Whereas teachers from some schools feel comfortable teaching their subjects, teachers from other schools do not appear prepared to teach their subjects. One teacher explains, “We do not charge any fees in this schools and teachers do not receive salaries or trainings. Among us there are teachers who taught for 20 years without having received any formal training.”<sup>277</sup> Another teacher admits that he/she does not have the training or materials to teach Somali: “Somali language is one of the most difficult subjects, and I feel that I'm not good at that subject. For example, Somali grammar is the most difficult and we don't have any books or dictionaries that we can teach from, so we didn't learn it ourselves.”<sup>278</sup> In some cases, teachers are forced to stop in the middle of their lessons and tell students to play outside while they try to research topics with which they are unfamiliar. Other tactics teachers use include searching for answers on Google and asking other teachers for advice.

However, some qualitative participants did note the direct effect the program has had on this issue: “We understood that mathematics is difficult for the students. We therefore shared this with CARE and CARE brought two teachers. There were also two teachers we substituted and replaced so that they could take part in further trainings.”<sup>279</sup> Additionally, teachers report using the following methods to test students’ learning periodically and to motivate students: (1) administering homework/classroom exercises, (2) using pictures, stories, drawing, and physical materials like sticks in helping students learn new topics, (3) using question and answer techniques during class, (4) moving students from the back of the classroom to the front, (5) repeating lessons for students, holding one-on-one lessons, and holding afternoon lessons, (6) encouraging parents to support their children in school, (7) holding competitions between students and administering awards for accomplishments, (8) holding monthly tests or tests at the end of chapters, (9) reading lessons out loud for students, and (10) verbally encouraging students.

### **Impact of CARE’s Numeracy training**

In addition to the assessment of changes in teaching quality from baseline to midline, this section also briefly investigates the impact of a specific SOMGEP-T teacher training intervention on teaching quality. As part of the project’s first year of implementation, CARE has conducted training in numeracy pedagogy with over 250 teachers (267 at the time of writing). We expect the training program to primarily have impacts on numeracy learning scores, but it is also possible that the training will improve pedagogy beyond strictly the teaching of mathematics.

To assess the impact of the training program on teaching quality, we analysed five overall indicators. We selected these indicators for two reasons: some of the indicators – such as the use of respectful language – are relatively easy to change, and may show rapid shifts in response to training. Others, such as the use of participatory methods, are core aspects of pedagogy that are likely to have been emphasized even in a class on numeracy-specific pedagogy.

The majority of the teachers observed at the midline in intervention schools, 23 out of 32 in total, had received recent numeracy training. However, as shown in the table below, the training program is not associated with major or consistent changes in teacher performance relative to untrained teachers in

---

<sup>276</sup> FGD – Teachers

<sup>277</sup> FGD – Teachers

<sup>278</sup> FGD – Teachers

<sup>279</sup> FGD – CEC

otherwise similar (i.e. intervention) schools. It is important to emphasize the small sample size utilized in this analysis, with just nine teachers observed in intervention schools at midline having *not* received the training. The direct effect of training, via comparisons to untrained teachers in the same set of schools, should be a priority of future evaluation waves.

*Table 67: Effect of teacher training on classroom demeanour, participation and the use of formative assessments*

Indicator	Baseline		Midline		
	Comparison	Intervention	Comparison	Intervention	
				Not Trained	Trained
Respectful language used			76.7%	66.7%	65.2%
Teachers use student names			53.3%	22.2%	52.2%
Teachers use formative assessment (self-reported)	59.3%	43.1%	73.3%	66.7%	73.9%
Teachers use formative assessment (as seen in records)	40.7%	26.2%	23.3%	44.4%	39.1%
Average of index for participation	0.56	0.57	0.64	0.71	0.68

SOMGEP-T seeks to improve the intermediate outcome of teaching quality by training teachers on (1) improving delivery of literacy and English language skills with the support of digital content, (2) improving delivery of numeracy, and (3) providing structured remedial support to students. The project’s design with regard teaching quality is evaluated to be GESI Accommodating. The intervention related to teaching quality, “acknowledges but works around gender, disability, or other social difference and inequality to achieve project objectives.” While the project’s Theory of Change acknowledges the marginalisation of girls, girls with disabilities, and barriers to better learning and transition outcomes, given that girls are approximately a grade or two behind, the teaching quality intervention is broadly aimed at improving the delivery of educational services for all students.

Toward the goal of improved educational services, some progress already appears to have been made despite the intervention only having been underway for less than a year. Active participation in classrooms have already improved since the baseline in intervention areas: 62 percent of teachers were observed asking students open-ended questions in the baseline and in the midline, 78 percent of teachers were observed doing so; similarly, in the baseline, 60 percent of teachers were observed asking students for their opinions in the baseline and in the midline, 75 percent of teachers did so. The self-reported use of formative assessments in intervention schools increased from 43.1 percent to 71.9 percent, and 26.2 percent of teachers were able to show records indicating that they used formative assessments while 40.6 percent could do so in the midline. The classroom environment also has improved in the midline with a lower rate of corporal punishment observed and reported by students, and a greater share of girls who say that their teachers make them feel welcomed in the midline, 77.2 percent, than in the baseline, 88.3 percent.

## 6.4. Life skills

Developing girls' life skills such as leadership skills, financial literacy, and business management are among the SOMGEP-T project's key intermediate outcomes. According to the project's Theory of Change, these skills correlate with girls' learning outcomes. This section presents an analysis of whether the SOMGEP-T project has had an impact on girls' leadership skills in the intervention group compared to the comparison group, using the difference-in-differences (DiD) approach. To examine girls' leadership skills, the indicators used here originate mainly from: 1) the CARE International Youth Leadership Index; and 2) the life skills module from the standard household survey template.

### *Youth Leadership Index (YLI)*

#### *YLI Questions:*

- q\_1 I like to try new activities that I may not know how to do.
- q\_2 My friends ask me for advice.
- q\_3 I recognize when people have different skills to contribute to a task.
- q\_4 I am comfortable when my teacher calls on me to answer a question.
- q\_5 I contribute ideas to discussions at home even if they are different from others' ideas.
- q\_6 I ask questions at school when I don't understand something.
- q\_7 I can describe my thoughts to others.
- q\_8 The things I do set a good example for my peers.
- q\_9 I consider possible outcomes of my decisions before making them.
- q\_10 I accept responsibility for the outcomes of my decisions.
- q\_11 I recognize when choices I make today can affect my life in the future.
- q\_12 I can show what is important to me with my actions.
- q\_13 If someone does not understand me, I try to find a different way of saying what is on my mind.
- q\_14 I encourage others to join together to help my community.
- q\_15 I cooperate with others to get things done at home.
- q\_16 If someone treats me unfairly at school, I am comfortable telling an adult.
- q\_17 I am willing to work hard to achieve my dreams.
- q\_18 I am better able to finish a task when I plan ahead.
- q\_19 When I have the opportunity, I can organize my peers to do an activity.
- q\_20 I am interested in being a leader at my school.
- q\_21 I try to understand the cause of a problem before trying to solve it.

Since the in-school girls are expected to participate in the girls' empowerment forums (GEFs), they were all asked to respond to the YLI questions. Together, 807 girls responded to YLI questions, and they were given a total of 21 questions with the options ranging from 1 to 4, with 1 indicating "Rarely," and 4 indicating "Almost always." To create the YLI scores, the YLI items were summed up, making 21 the lowest possible score and 84 the highest possible score.

The histogram below presents the distribution of girls' YLI scores, showing a slightly right skewed distribution. The Cronbach's alpha of YLI score is 0.91, indicating a very high level of internal consistency.

It appears that girls' average YLI score has increased from 53.3 in the baseline to 54.9 at the aggregate level in the midline, showing a percentage difference of 2.95.<sup>280</sup> This difference is below what the project aimed to achieve in the midline (5 percent above baseline). The result of difference-in-differences shows no significant differences in YLI scores between the intervention and comparison groups over time. The girls in the comparison group received a slightly higher average YLI score from the baseline (52.7) to midline (55.1) with 4.45 percent increase compared to the intervention group whose girls had 1.6 percent increase from the baseline (53.9) to midline (54.8).

When data was disaggregated by age, grade, disability and zone, the girls' YLI score did not vary significantly across the subgroups either. The results of difference-in-differences analysis for all girls, girls 12 years of age or older, girls with any kind of disability, and girls from each zone are indicated in Table 68.

Figure 30: Histogram of YLI Scores (in-school girls)

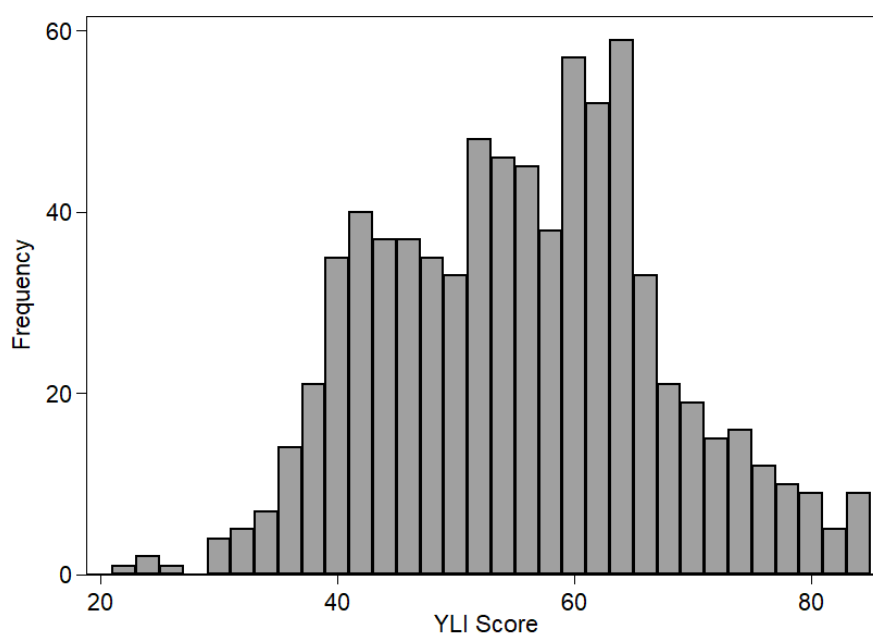


Table 68: Project impact (difference-in-differences) on self-confidence

Indicators	Difference-in-Difference	Age	Grade	Any
		(>=12)		disability
YLI Score (n=807) <sup>281</sup>	-1.43	-2.46	-1.4	-2.1

<sup>280</sup> The skewness and kurtosis for the midline data were 0.08 and 2.61, respectively.

<sup>281</sup> The sample size only includes the sample in the midline.

The information provided in the table below on girls' average score over time among the intervention and comparison group, explains why the difference-in-difference has a negative coefficient (-1.43). On average, the girls in the intervention group achieved a smaller increase (0.9) in their YLI score from baseline to midline compared to their counterpart in the comparison group (2.41), resulting to a negative average difference-in-difference score of -1.51.

Moreover, the percentage of GEF participation among the girls with 12 years of age and above and enrolled in school, have overall increased from 16.0 percent in the baseline to 33.4 percent in the midline, yet GEF participation doesn't seem to be a strong predictor of girls' YLI score in the midline.<sup>282</sup>

*Table 69. In-school girls' average YLI score*

Indicator	Baseline Intervention Schools	Midline Intervention Schools	Intervention Diff	Baseline Comparison Schools	Midline Comparison Schools	Comparison Diff	Diff in diff
YLI score	53.9	54.8	0.9	52.7	55.11	2.41	-1.51

### *Testing the ToC-YLI score*

#### **Learning outcomes**

The project's theory of change proposes that there is a positive relationship between girls' life skills and their learning outcomes, meaning the more leadership skills, self-confidence, and self-efficacy girls acquire, the better performance they are expected to have in school as a result of more confidence and abilities to participate in the classroom.

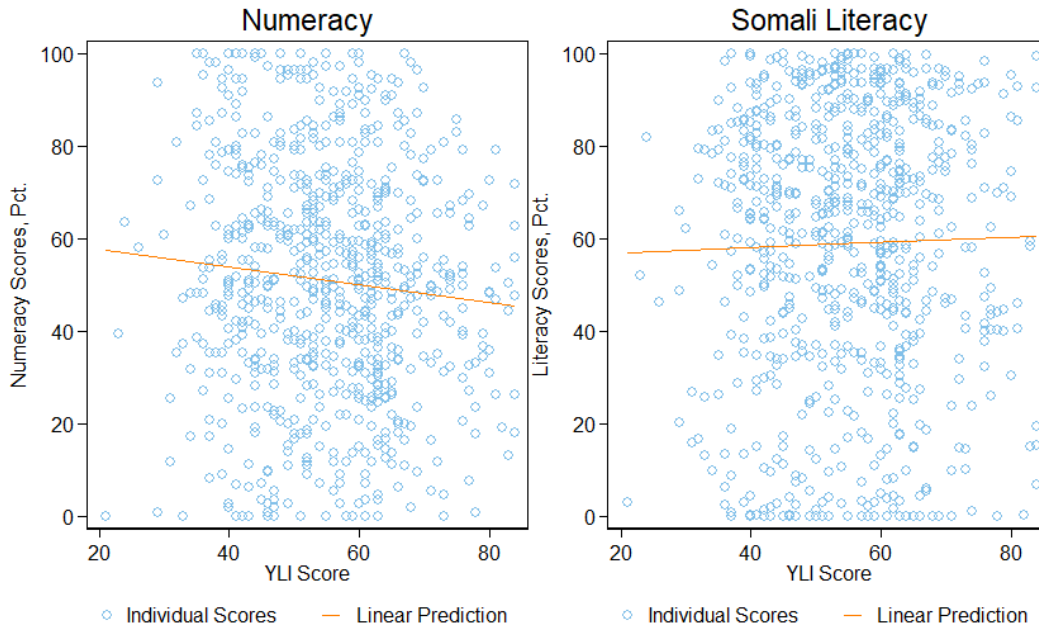
When the YLI scores were regressed against their learning outcomes in the midline, they were found to be negatively correlated with the numeracy outcomes,<sup>283</sup> but positively correlated with the Somali literacy outcomes.<sup>284</sup> The relationship with numeracy outcomes is statistically significant, while the relationship with Somali literacy is not significant (Figure 31).

<sup>282</sup> In a regression of YLI score predicted by GEF participation, it has a coefficient of -0.77 and p-value > 0.05

<sup>283</sup> YLI score and numeracy outcomes has a coefficient of -0.21 and p-value of < 0.05.

<sup>284</sup> YLI score and Somali literacy outcomes has a coefficient of 0.03 and p-value of > 0.05.

Figure 31. YLI scores and learning outcomes



### Attendance

In addition to learning outcomes, attendance has been hypothesized in the ToC to be positively correlated with leadership skills. Girls are expected to have greater levels of attendance and participation in school because of the strengthened leadership skills they gain from the participation in GEF activities. The result of regression of caretaker-estimated attendance against YLI scores in the midline data indicates however that leadership skills is not a significant predictor of attendance.<sup>285</sup>

<sup>285</sup> YLI score and attendance has a coefficient of -0.04 and p-value of >0.05.



## Life Skills Module

The project's ToC suggests that if girls improve their life skills, they will greater engagement in school as a result of more confidence and participation in the classroom. This section provides an analysis of girls' leadership skills, as well as their self-esteem and sense of agency. The life skills questions are divided into three categories: learning to learn, learning for life, and agency.<sup>286</sup> The table below lists the full set of questions for reference.

Table 70: Life-Skills Survey Questions

	Questions	Respondents
Learning to learn	I can read as well as my friends	Everyone
	I am as good at math as my friends	
	I get nervous when I have to speak in front of an adult	Everyone excluding in-school girls <12
	I get nervous when I have to speak in front of a group of people my age	
	I get nervous when I have to read in front of others	In-school girls
	I get nervous when I have to do math in front of others	
	I feel confident answering questions in class	Everyone excluding in-school girls <12
	I feel confident answering questions when I'm in a group of people	
Learning for Life	I would like to continue studying/ attending school after this year	In-school girls
	I would like to continue learning by going back to school, learning a vocation or trade	Out-of-school girls
	I recognize when choices I make today about my studies can affect my life in the future	Only in-school girls >=12
	I recognize when choices I make today can affect my life in the future	Only out-of-school girls >=12
	I can describe my thoughts to others when I speak	Everyone
	I can work well in a group with other people	
	When I have the opportunity, I can organize my peers or friends to do an activity	
	I often feel lonely	Only out-of-school girls >=12
	I have trusted friends I can talk to when I need to	Girls >=12
	I have trusted adults I can talk to when I need to	
	I ask an adult if I don't understand something	Out-of-school girls
	I ask the teacher if I don't understand something	In-school girls
	Agency	Who decides: Whether or not you will go to school
Who decides: Whether or not you can go back to school or vocational training		Out-of-school girls
Who decides: Whether or not you will continue in school past this grade		In-school girls
Who decides: If you will work after you finish your studies		
Who decides: How often you spend time with your friends		Everyone
Who decides: When/ at what age you will get married		

<sup>286</sup> DFID underlined these categories and questions as the life skills indicators of concern.

Who decides: Who decides what type of work you will do after you finish your studies	
Who decides: How you spend your free time	Girls >=12

These questions are analysed mainly through the examination of key indicators of life skills attainment as well as standardization of the life skills index. The presentation of results is disaggregated by intervention versus comparison, age, and school status. Some of these questions are common across groups while others were designed exclusively to the characteristic of each group. The number of questions ranges between 15 and 23 and is generally divided into four sets for the sampled girls, depending on their school status (in school versus out of school) and age (<12 or ≥12). The questions related to learning and transition have Likert-type responses ranking between “strongly agree” with a score of 1 to “strongly disagree” with a score of 5, while questions on agency provide “I decide,” “I decide jointly with my family,” or “my family decides” as response options. The percentages of the girls stating “I decide” and “I decide with my family” and those saying “strongly agree” and “agree” are illustrated in the tables below.

*Table 71: Reported as percentage stating “I decide” or “I and my family decide jointly”*

Summary table		Agency							
		Whether or not you will go/back to school	Whether or not you can go back to school or vocational training	Whether or not you will continue in school past this year	When/ at what age you will get married	If you will work after you finish your studies	What type of work you will do after you finish your studies	How you spend your free time	How often you spend time with your friends
Intervention	I decide	31%	58%	31%	33%	46%	47%	47%	46%
	Decide jointly	40%	26%	42%	41%	37%	40%	40%	39%
Comparison	I decide	24%	42%	26%	34%	33%	42%	38%	36%
	Decide jointly	44%	39%	46%	40%	45%	41%	43%	42%
Under 12's	I decide	23%	20%	22%	26%	33%	31%	0%	29%
	Decide jointly	40%	60%	44%	39%	38%	46%	75%	45%
12 and over	I decide	29%	53%	30%	35%	42%	47%	43%	44%
	Decide jointly	42%	30%	44%	41%	41%	39%	41%	40%
In school girls	I decide	28%	50%	29%	33%	40%	44%	42%	41%
	Decide jointly	42%	50%	44%	41%	41%	41%	42%	41%
Out of school girls	I decide	20%	49%		46%		54%	58%	54%
	Decide jointly	60%	32%		44%		39%	36%	42%
<b>Sample size (valid responses)</b>		767	45	762	807	762	807	673	807

Table 72: Reported as percentage stating “strongly agree” & “agree”

Summary table	Learning to learn							
	I can read as well as my friends	I am as good at math as my friends	I get nervous when I have to speak in front of an adult	I get nervous when I have to speak in front of a group of people my age	I get nervous when I have to read in front of others	I get nervous when I have to do math in front of others	I feel confident answering questions in class	I feel confident answering questions when I'm in a group of people
Intervention	98%	94%	38%	25%	36%	34%	91%	86%
Comparison	98%	93%	39%	29%	35%	34%	89%	88%
Under 12's	98%	93%	56%	33%	30%	35%	92%	78%
12 and over	98%	94%	38%	27%	37%	34%	90%	87%
In school girls	98%	94%	37%	26%	36%	34%	90%	87%
Out of school girls	88%	78% <sup>287</sup>	54%	46%				90%
<b>Sample size (valid responses)</b>	<b>807</b>	<b>807</b>	<b>678</b>	<b>678</b>	<b>762</b>	<b>762</b>	<b>762</b>	<b>678</b>

For indicators relating to girls' decision-making power or agency, particularly when it comes to deciding whether to go back to school, continuing in school past this year, and when to get married, the girls tend to be less independent in making decisions by themselves. Yet, the girls in the intervention group and those above 12 have more frequently stated that they can decide about more personal issues such as how to spend their free time and time with friends. Concerning indicators that relate to confidence and self-esteem, the girls have exhibited high scores in most of the cases. For example, most of the girls believe that their reading and math skills are as good as their friends' skills. Except for the out-of-school girls and those under 12, the minority of girls feel nervous when they have to speak in front of an adult and in a group of people or when they have to read and do math in front of others. As shown in Table 72: Reported as percentage stating “strongly agree” & “agree”, a large majority of girls also feel confident in terms of answering questions in class or in a group of people. They also seem to feel competent in describing their thoughts to others, working in a group of people, and organizing peers to do an activity which is an indication of strong leadership and organization skills among girls. The quantitative and qualitative findings, however, suggest that lack of peers' support and shyness among girls hinder their participation in the classroom. As indicated in the table below, this may be a more serious issue for the girls with disabilities as they may experience higher levels of loneliness in the school.

<sup>287</sup> We believe that asking the out-of-school girls about their reading and math skills is relevant because some of them were in school previously and therefore, can read and do math.

Table 73: Reported as percentage stating “strongly agree” & “agree”

Summary table	Learning for life (Transition)											
	I would like to continue studying / attending school after this year	I would like to continue learning by going back to school, learning a vocation or trade	I recognize when choices I make today about my studies can affect my life in the future	I recognize when choices I make today can affect my life in the future	I can describe my thoughts to others when I speak	I can work well in a group with other people	When I have the opportunity, I can organize my peers or friends to do an activity	I often feel lonely at school	I have trusted friends I can talk to when I need to	I have trusted adults I can talk to when I need to	I ask an adult if I don't understand something	I ask the teacher if I don't understand something
Intervention	96%	84%	95%	83%	94%	94%	91%	33%	95%	96%	90%	96%
Comparison/Control	97%	100%	95%	95%	92%	93%	90%	41%	93%	95%	96%	95%
Under 12's	98%	100%	100%		91%	92%	88%		100%	100%	80%	96%
12 and over	97%	93%	95%	90%	94%	94%	91%	38%	94%	95%	95%	96%
In school girls	97%	75%	95%	75%	93%	94%	91%	25%	95%	96%	75%	96%
Out of school girls		95%		92%	90%	88%	85%		89%	92%	95%	
<b>Sample size (valid responses)</b>	<b>762</b>	<b>45</b>	<b>633</b>	<b>40</b>	<b>807</b>	<b>807</b>	<b>807</b>	<b>40</b>	<b>673</b>	<b>673</b>	<b>45</b>	<b>762</b>

Table 74: Reported as percentage stating “strongly agree” and “agree” (for girls with disability)

<b>Summary table</b>	
	I often feel lonely
Intervention	60%
Comparison/Control	40%
Under 12’s	-
12 and over	50%
In school girls	-
Out of school girls	44%
<b>Sample size (valid responses)</b>	10

### Life Skills Indicators

Girls with higher levels of confidence and participation in the classroom may perform better in school. Therefore, in the indicator method, only three indicators have been selected to assess girls’ levels of self-confidence with respect to their ability to speak in front of others and answer questions in class or in a group of people. These indicators are the same indicators that were used in the index method. Next, the results are analysed by subgroups to see if girls’ scores differ by school attendance, zone, and GEF membership.

As indicated in Table 72, although most of the girls believe that their reading and math skills are as good as their friends’ skills, they seem to feel less confident when it comes to speaking in front of an adult, in a group of people, and having to read and do math in class. Overall, the majority of girls rated themselves capable of speaking in front of an adult (53.3 percent) or in front of a group of people at their age (66.5 percent). They also (87 percent) reported they feel confident answering questions when they are in a group of people. Reading and math skills seem to be strong among the in-school girls because more than half of them did not believe that having to read (56 percent) or to do math (58.4 percent) in class makes them nervous. Most of them (90 percent) were also positive about their ability to answer questions in class.

In general, the SOMGEP-T project appears to have enhanced girls’ levels of self-confidence among the intervention group across some of the selected indicators. Yet, this difference between the intervention and comparison group is not statistically significant. More specifically, girls in the intervention group tend to be more capable of speaking in front of an adult (55 percent) or a group of people at their age (67.2 percent) compared to their counterparts in the comparison group (51 percent and 65.5 percent, respectively). However, there are fewer girls in the intervention group (86.1 percent) than in the comparison group (87.9 percent) who reported themselves confident answering questions in a group of people. Even though, girls in the intervention group rated their reading skills (55.8 percent) marginally lower than the girls in the comparison group (56.2 percent), they appear to have more confidence in their math skills (59.1 percent) and answering questions in class (90.5 percent) compared to the girls in the comparison group (57.3 percent and 89 percent, respectively). These percentages are demonstrated in the figures below.

Figure 32. Midline girls' responses (by intervention)

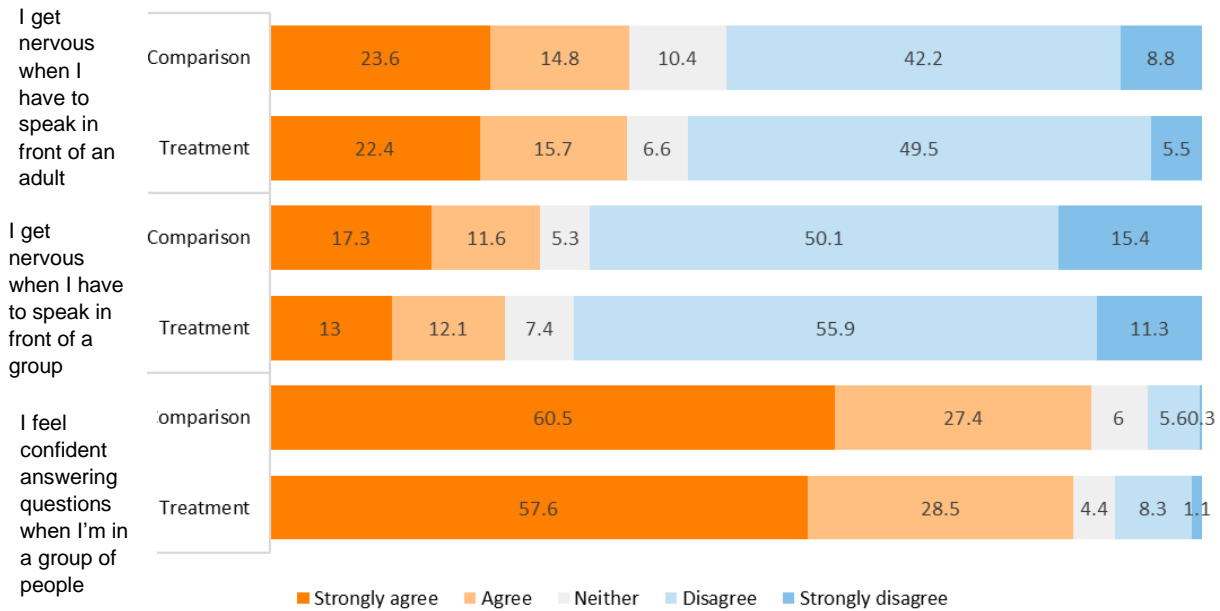
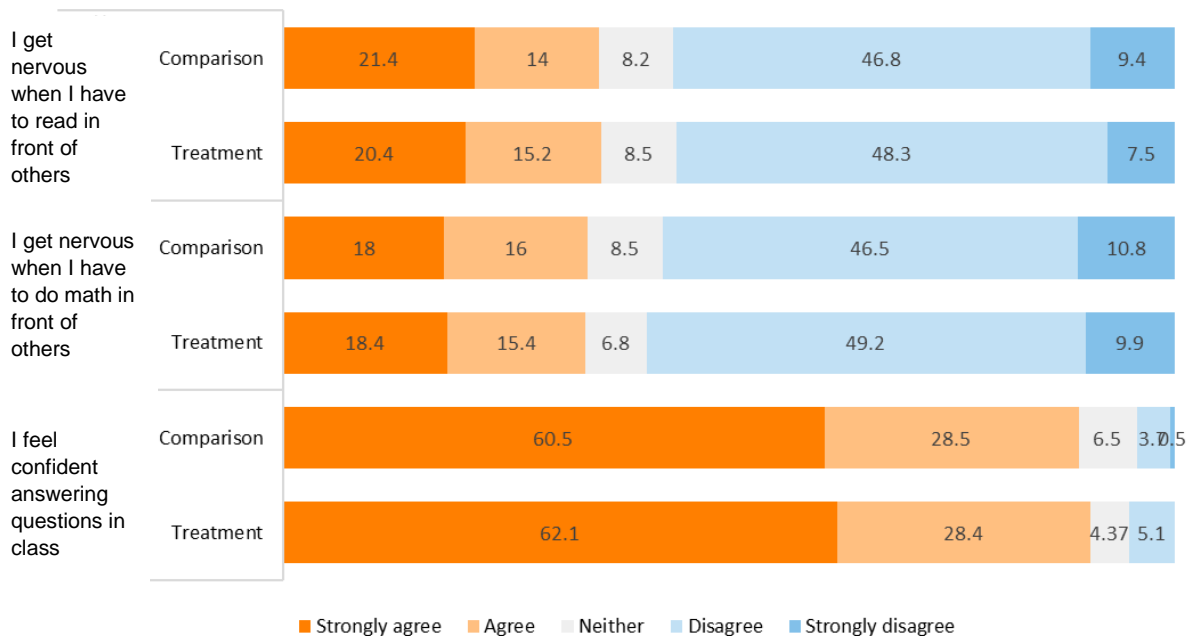


Figure 33: Midline girls' responses (by intervention)



When girls' responses are further disaggregated by school status (in school or out of school) and age, neither girls with 12 years of age and above nor those who were enrolled significantly excelled in their ability to speak, read, do math, or answer questions in class or in a group of people, compared to their counterparts. However, the girls who had any kind of disability seem to have improved their confidence answering questions in a group of people by participating in the SOMGEP project, but their abilities differ significantly across other indicators. No significant impact was either found across grades. The results of difference-in-differences are reported in the table below.

*Table 75: Project impact (difference-in-differences) on self-confidence*

Indicators	Difference-in-Difference				
		Enrolment (In-school)	Age (>=12)	Grade	Any disability
I get nervous when I have to speak in front of an adult (n=766) <sup>288</sup>	0.15	0.12	0.12	-	-0.05
I get nervous when I have to speak in front of a group of people my age (n=766)	0.036	0.003	-0.14	-	-0.73
I feel confident answering questions when I'm in a group of people (n=678)	0.20	0.16	0.29	-	1.1*
I get nervous when I have to read in front of others (n=762)	0.09	0.09	0.05	0.78	0.7
I get nervous when I have to do math in front of others (n=762)	0.03	0.03	-0.006	0.49	0.4
I feel confident answering questions in class (762)	-0.14	-0.14	-0.18	0.3	0.3

In the midline, girls with GEF membership did not also appear to have significantly different levels of reading<sup>289</sup> and math skills<sup>290</sup> or confidence answering questions in class compared to the girls in the comparison group who did not have membership.<sup>291</sup>

There was one life skills qualitative indicator included in the project logframe: Girls feel comfortable expressing themselves at school, in the community, and at home. In the midline, the target for this indicator was "Girls describe examples of engagement with others at school and household to express their needs and aspirations." Data collected from the FGDs with mothers and teachers and KIs with girls with disabilities were used to inform the analysis for this indicator. The qualitative evidence suggests that girls' participation in decision-making and engagement in their homes, schools, and communities is improving, but some barriers to girls' engagement do remain. There was a lack of consensus among mothers both within specific FGDs and across FGDs on whether girls feel comfortable expressing their ideas and aspirations at home, in school, and in their communities. The majority of mothers appear to feel girls can share their ideas in all three of these areas, but there are often caveats.

<sup>288</sup> The sample sizes shown in this table only include the sample in the midline.

<sup>289</sup> In a regression of reading skill score predicted by GEF participation, p-value > 0.05

<sup>290</sup> In a regression of math skill predicted by GEF participation, p-value > 0.05

<sup>291</sup> In a regression of confidence answering questions in class predicted by GEF participation, p-value > 0.05



For example, some mothers explained that girls express small concerns, but do not necessarily participate in “big” decisions: “Children can express their small concerns/Ideas but not ideas about the big issues.” Others explained that girls will share their ideas and opinions if they believe their mothers will support them, but will not share their ideas if they feel their mothers will disapprove of them: “Yes, we accept many decisions she suggests but if we do not take her decision, she will not tell us anything and will no longer want to take part decision making.” Another mother adds, “They share with me their issues they know that I will accept and agree with, but other issues they will not share with me.” These appear to be typical parent/child interactions, but in other cases, mothers’ responses suggested that girls’ views are not always valued because of their age: “We always decide for them because they are children and they don’t know what is good for them.” These findings are consistent with the findings of the baseline, during which mothers shared that girls themselves are often the ones who make the decision to drop out of school, get married early, or spend time with boys their parents would disapprove of.

Girls’ shyness at home, in school, and in their communities was a theme that emerged across groups, which is also consistent with the findings from the baseline. When asked whether girls and boys participate equally in class, teachers most commonly reported that girls are shy and that boys are more confident/active in class. In some cases, this appears to be due to the fact that girls have more difficulty preparing for lessons outside of school, as they are busy with household chores. As one teacher explains, “Many girls have to do house chores after school which affects their ability to learn (do homework or study after class) and to pass to the next grade.” In other cases, it appears that girls’ shyness is a product of their culture/environment. One teacher explains, “Girls are always like this. There is no interaction, and this is just the culture.” Another shares that girls are not just shy in terms of participating in class: “No, girls are even shy when they are just sitting in their chair.” Mothers also provided evidence that girls are shy at home, particularly when they are young. When asked whether her daughter shares her ideas/decisions, one mother explains, “No she does not share with me her decisions because she is young, and she feels shy. On the other hand, a 15-year-old girl illegally migrated abroad, and she did not tell her family.”

There is also evidence, however, that girls are becoming more confident in sharing their ideas. One mother explains how girls now push back if their parents decide not to send them to school but send the boys in the family: “If you do not enrol the girl at the school, she will say to you why you did you not admit me to the school? So you either admit both the boy and the girl, or do not enrol either of them.” Additionally, although most teachers expressed that girls are very shy in class, many also expressed that girls are more committed to their school work than boys and that the situation has changed with time. As one teacher explains, “Most of the time girls are shy, but that does not mean that there is a lack of knowledge. There was a time when boys were more active than the girls, but right now everything has changed. We can now have competitions between boys and girls [thanks to] social media, mobile phones, and TV. These competitions have become helpful.”

Girls themselves also shared that they feel comfortable speaking up in their homes, schools, and communities, and despite the discrimination girls with disabilities face, in many cases, they appear to be receiving support from their parents, fellow students, and community members in completing their education. When asked whether her parents support her in her schooling, one girl with disability explained, “They support me very well, for example at night my father reads me the lesson and sometimes he helps with examination and he encourage me with the way I read the lesson.” There is also evidence that other students provide girls with disabilities with assistance when they are struggling, suggesting that these girls feel comfortable either asking for or accepting help from their fellow students: “Sometimes I am sick in class - for example, my rip-cage is painful and I cannot write anything. Then, other students in the class helped me write the lesson.” These results suggest that the midline target for the life skills indicator has been reached, but that there is still work to be done to encourage girls to constructively engage with their parents and in class.

## Life Skills Index

In this section, a standardization method was adopted to create the index scores where each girl received a life skills score between 0 and 1, by combining the indicators listed in the first table presented in this section. Next, the results are disaggregated by age, school status, zone, and GEF membership to see the effects across subgroups.

The life skills index for each group of girls have been constructed separately, because they received a slightly different set of questions depending on their age and school status. It is worth reiterating that limitations in the midline did not allow the evaluation to collect data from the baseline out-of-school girls, and therefore, all girls surveyed in the midline attended school last year in the baseline. There was a total of 43 girls 12 years of age and above who were in school during baseline but dropped out of school by the time of the midline. While these girls were included in the midline analysis and are considered to be out-of-school, the indicators used to construct the life skills index for these girls included only those that were asked to respond in both the baseline and midline. There were no girls younger than 12 in the midline who were in school during the baseline, and as such, they are not discussed in the following analysis.

The results of the difference-in-differences analysis across the girls' subgroups are shown in Table 76 below and reveals that all the girls in the intervention group in the midline generally have better life skills scores compared to the girls in the comparison group as a result of participation in the SOMGEP-T program, but this difference is not statistically significant. To show the extent of improvement in the life skill index scores over time between the intervention and comparison groups, the average life skill score for each group of girls is reported in Table 77. Unlike the comparison group, girls in the intervention group, on average, have received higher life skills score from baseline to midline, resulting to a positive average difference-in-difference index score.

It is worthwhile mentioning that the project aimed to achieve 5 percent increase in the average score from the baseline to midline. At the aggregate level, the out-of-school girls with 12 years of age and above had a percentage difference of 20 in the average life skills index score from the baseline (0.54) to midline (0.66). Looking at the data by intervention, it appears that the intervention group received a 45 percent increase in average life skills index score from the baseline to midline while the comparison group had a 1.47 percent decrease in its average life skills index score during the same period. Similarly, the older in-school girls scored higher in the midline compared to their baseline average score with a 2.7 percent difference. The girls in the intervention group seem to have 5.5 percent increase in their midline average score compared to the girls in the comparison group who did not have any difference in their average score from the baseline to midline. Additionally, the younger in-school girls showed also some improvement in their midline average score with 11.3 percent difference from the baseline. The intervention group of younger in-school girls had slightly higher percentage increase (13.5) in their midline average score compared to their counterparts in the comparison group (10.5 percent).

When data was further analysed by girls' disability status, grade and zone, no significant project impact was observed in the life skill scores across the subgroups of girls. Though insignificant, the girls with 12 years of age and older and enrolled in school have scored higher compared to their counterparts. This increase in life skill scores can be observed across the three regions. The results of the difference-in-differences analysis is illustrated in the table below.

*Table 76: Life Skills Index Difference in Difference Results*



	Difference-in-Difference	Any disability	Grade
Out-of-school girls >=12 (n=36) <sup>292</sup>	0.21	-	-
In-school girls >=12 (n=629)	0.04	-	0.04
In-school girls <12 (n=132)	0.01	-0.06	0.01

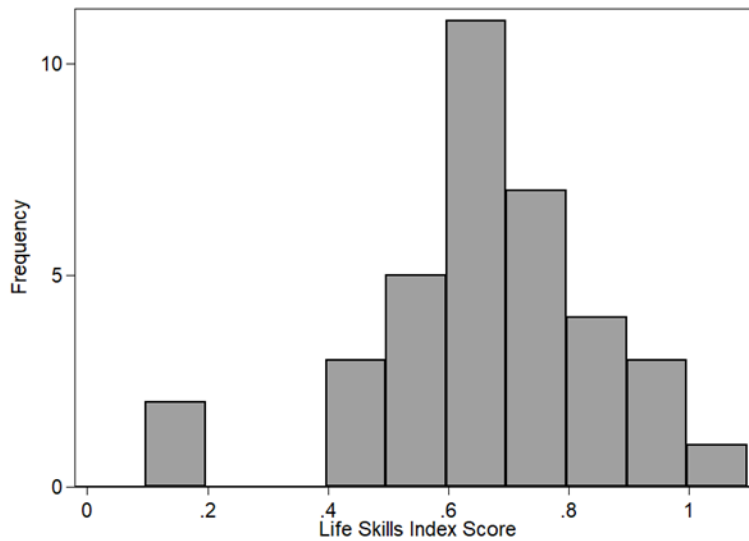
*Table 77: Average Life Skills Index Score*

Indicator	Baseline Intervention Schools	Midline Intervention Schools	Intervention Diff	Baseline Comparison Schools	Midline Comparison Schools	Comparison Diff	Diff in diff
Out-of-school girls >=12	0.44	0.64	0.2	0.68	0.67	-0.01	0.21
In-school girls >=12	0.7	0.74	0.4	0.72	0.72	0	0.4
In-school girls <12	0.59	0.67	0.08	0.57	0.63	0.06	0.02

The distributions of these scores in the midline are summarized below for each subpopulation, with a brief description of the properties of the score.

<sup>292</sup> The sample sizes shown in this table only present the sample size in the midline.

Figure 34: Index Score for Out-of-School Girls Over or Equal to 12 (Midline)

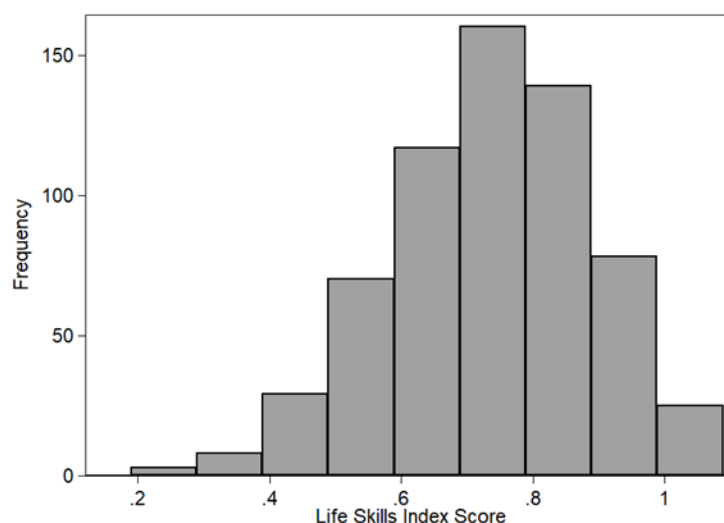


The average score for out-of-school girls over or equal to 12 is 0.66 in the midline, with a minimum score of 0.096 and maximum of 0.98. Half of these girls achieved scores between 0.54 and 0.78.<sup>293</sup> The Cronbach's alpha of the out-of-school girls' score is 0.761. Depending on the standards employed, the Cronbach's alpha scores reported here may be slightly below the level desired. For instance, a common rule of thumb is that alpha should be above 0.8 for an index. In the context of a psychological measure as diffuse and wide-ranging as life skills, we should expect and accept lower alpha scores.<sup>294</sup> In addition, formative research with scales and items that have not been extensively validated -- such as those used in this section -- also suggest accepting a lower score. For these reasons, we view the scores reported, all above 0.7, to indicate indices of reasonable consistency.

<sup>293</sup> The skewness and Kurtosis are -0.75 and 4.07, respectively.

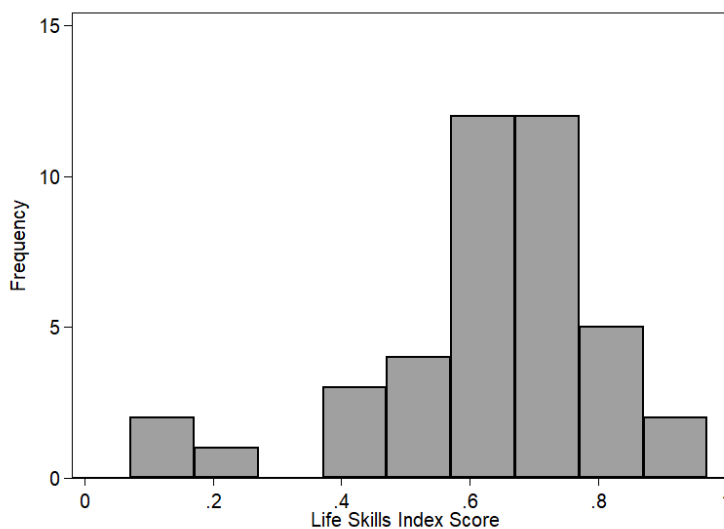
<sup>294</sup> Lance, C., Butts, M., & Michels, L. (2006). The Sources of Four Commonly Reported Cutoff Criteria: What Did They Really Say? *Organizational Research Methods*, 9(2), 202–220. <https://doi.org/10.1177/1094428105284919>

Figure 35: Index Score for In-School Girls Over or Equal to 12 (Midline)



Similarly, the average score of the in-school girls over or equal to 12 has increased from 0.71 in baseline to 0.73 in midline, with the lowest score being 0.18 and the highest of 1. Moreover, 50 percent of girls' scores ranked between 0.63 and 0.83.<sup>295</sup> The Cronbach's alpha for this group of girls' score is 0.78. When data is disaggregated by girl's participation in GEF activities in the midline, it appears that the girls' life skills index score in the intervention group is significantly higher for the girls who participated in GEF activities compared to those in the comparison group who did not.<sup>296</sup>

Figure 36: Index Score for In-School Girls below 12 (Midline)



<sup>295</sup> The skewness and Kurtosis are -0.39 and 2.85, respectively.

<sup>296</sup> In a regression of life skills index score predicted by GEF participation, the coefficient was 0.04 and p-value < 0.05.

Similar to the older girls, the younger in-school girls' midline average score (0.65) has slightly improved compared to the baseline average score (0.58). The scores of these girls range between 0 and 1, and half of the girls' scores fall between 0.57 and 0.77.<sup>297</sup> The Cronbach's alpha for the life skills score of the younger in-school girls is reported to be 0.73.

To summarise the findings, it was found that the SOMGEPT project has slightly improved the girls' leadership and life skills. The analysis showed that while the girls in the intervention group gained smaller increase in the average YLI score compared to that of the girls in the comparison group, they have (excluding the younger out-of-school girls) more confidence in their life skills by scoring higher than the comparison group. More specifically, girls in the intervention group generally feel more confident speaking in front of an adult or in a group of people as well as doing math and answering questions in class. However, none of these differences between the intervention and comparison group were statistically significant. The only subgroup girls who had significant difference in difference in their level of confidence speaking in front of a group was the girls with disabilities.

## **6.5. Community-based attitudes and behaviour change**

In the baseline, information gathered through the qualitative portion of the research was used to understand and illustrate community attitudes toward education generally and girls' education in particular. Broadly, the results suggested that education is highly valued by community members across Somalia and is seen as the means through which individuals can support themselves and their families, improve their communities, and even benefit the national and global economy. However, the findings from the qualitative data suggested that there is still somewhat of an active debate in Somali society on the importance of girls' education in relation to boys' education. On the positive side, qualitative interview participants, particularly mothers, highlighted how there has been a major shift in community attitudes toward girls' education in the last generation – in the past, girls' education was not valued or, as a result, prioritized, and therefore women did not have much opportunity to work outside of the home. This positive shift in cultural attitudes was perhaps most evident among girl and boy students, as both groups seemed to accept without question the idea that education is important and that women and men can and should hold similar roles after they complete their schooling – as teachers, doctors, NGO workers, politicians, or any other position they aspire to hold.

Although most mothers, teachers, CEC members, and students appeared to have positive attitudes toward girls' education, others explicitly stated that boys' education is more important than girls' education, and it was evident from some discussions that girls still face barriers to enrolling in and attending school that boys do not face. The respondents who seemed to feel boys' education is more important than girls' education described boys as 'leaders' and 'more useful' because, according to these respondents, boys assume more responsibility than girls when they become adults. Others described the unique barriers that prevent girls from enrolling in, attending, and succeeding in school. These barriers include prioritization of boys' education over girls' education for families facing financial constraints, prioritization of boys' education due to the expectation that girls will get married early, the use of girls for household chores, girls' absence during menstruation due to feelings of shame, and low morale among girl students. A number of these factors indicated that there are still social norms that restrict girls' access to school and that societal attitudes toward girls and girls' education have been internalized by girls in ways that negatively affect their success in school.

For the midline evaluation, the results from the qualitative portion of the research were again analyzed in this section of the report, but this time with particular attention paid to differences in community attitudes since the baseline evaluation. There were, however, some changes made to the methodology from

---

<sup>297</sup> The skewness and Kurtosis are -0.45 and 4.0, respectively

baseline to midline that are worth noting to provide context for findings highlighted hereafter. In the midline, the FGDs with girl and boy students were eliminated. Instead, the following three qualitative elements were added to the research: 1) Key informant interviews with girls with disabilities; 2) A participatory risk mapping exercise; and 3) A participatory story-telling exercise. The risk mapping exercise will be discussed in more detail in the following section on school-related and gender-based violence, but the analysis in this section relies in part on the results from KIIs with girls with disabilities and the story-telling exercise. Because these two elements are new to the midline, we have provided below a brief description of the composition and purpose of each element.

In the key informant interviews with girls with disabilities, the girls were asked a myriad of questions aimed at understanding the unique experiences of girls with disabilities in school and in their communities. Girls were asked to describe their experiences traveling to and from school, experiences with other students and teachers, causes of absence, experiences learning in the classroom and carrying out classroom activities, and the attitudes of friends, family, and community members toward their education. They were also asked to detail the treatment and support they have received as a result of their disability, the impact their disability has had on their relationships and level of participation in school and community activities, and the aspirations and hopes they have for the future. The findings from these interviews allowed us to deepen the analysis in this section, adding an extra layer of first-hand experiences from some of the most marginalized members of Somali society.

In the story-telling/vignette exercise, girls were presented with a number of different short stories about girls in Somali society and asked to discuss among themselves how each story might end. The indirect questioning used in this story-telling exercise allowed us to control for the social desirability bias we might have seen if we had asked girls to share their experiences directly. Girls were instead encouraged through this method to project their own feelings and perceptions onto a hypothetical third-party, which strengthens the likelihood that their responses provide us with an accurate representation of their communities' attitudes toward girls' education.

The findings from these two new elements were analysed alongside the findings from the FGDs with CEC members, mothers, and teachers, and KIIs with REOs for this section of the report. Additionally, we present a number of quantitative indicators and analyse changes in these indicators from baseline to midline. Taken as a whole, the results suggest that community attitudes toward girls' education have improved since the baseline. The table below analyses agreement with two statements, based on round and intervention status (i.e. difference-in-differences): (1) "Even when funds are limited it is worth investing in [girl name]'s education" and (2) "A girl is just as likely to use her education as a boy." For each, we have classified caregivers as agreeing if they agreed strongly with the statement. The results in the table are the share of caregivers in each group that agreed.

There has been a notable, positive change in the intervention group – for both indicators, the proportion of caregivers who agree with the statement has increased approximately 10 percentage points (79.8% to 89.1% and 72.2% to 82.1%). For the comparison group, the proportion who agree with the first statement actually decreased slightly (83.0% to 82.2%) and there was a less of an increase than was observed in the intervention group for the second statement (75.8% to 81.9%). As indicated by the asterisk, the difference-in-differences estimate of project impact is statistically significant for the first statement (a 10.1 point increase in intervention communities, relative to comparison communities, from baseline to midline). Notably, it is also significant if we limit the analysis to "panel girls" (i.e. girls where we interviewed the exact same set of girls from baseline to midline).



Table 78: Caregivers' perceptions on worth of girls' education

Outcome	Intervention Baseline	Comparison Baseline	Intervention Midline	Comparison Midline	Difference-in-Differences
Girls' education is worth investing in	79.8	83.0	89.1	82.2	10.1*
A girl is just as likely to use her education as a boy	72.2	75.8	82.1	81.9	3.9

These quantitative findings are supported by the findings in the qualitative data. Almost all midline qualitative interview participants appear to feel girls' education is just as important as boys' education. Whereas in the baseline there was some debate within focus groups as to the relative importance of girls versus boys, in the midline, there is only one example of a mother explicitly voicing discriminatory attitudes toward girls: "I believe boys are better than the girls in terms of the education because men are wiser than the women, although it depends on the effort the person makes. I believe men are always better than women."<sup>298</sup> Otherwise, there now appears to be a firm consensus around the importance of girls' education. As has been noted in previous sections, there was even evidence in the midline that pastoralist families are now invested in their girls' education after losing their livestock. As one REO explains, "I used to have parents say to us that the livestock will run away if their children do not supervise them as they did not understand the importance of the education. Parents also use to say look at the university graduates that do not have any jobs. But since the droughts have happened and killed most of their livestock, parents have brought their children to schools."<sup>299</sup>

Additionally, almost all of the mothers interviewed in the FGDs expressed that their girls wish to continue their education after they complete their current levels and also expressed their own personal support for these education-oriented goals. The majority of interviewed girls with disabilities also reported that their future goals include continuing their education, with many expressing that they would like to continue up to the university level. As one girl explains, "I would like to continue my education up to the university level and then find a job, because I want to help my parents and my coming children."<sup>300</sup> Financial constraints appear to be the only factor that discourages mothers from continuing to send their daughters to school and girls with disabilities from aspiring to continue their education. When one mother was asked whether she will be able to continue sending her daughters to school after primary school, she replies, "No, we are already struggling a lot with primary education because we cannot afford books and uniforms."<sup>301</sup> She further explains that "there is no secondary school in this community - only in Burao or big cities. Only wealthy families can afford that."

Despite these positive findings, qualitative interview participants across groups did still mention early marriage, pregnancy, prioritization of boys' education over girls' education, and absence during menstruation as unique barriers girls face to enrolling in, attending, staying in, or succeeding in school. The quantitative data also holds some evidence that girls face additional barriers that boys do not face. The graph below presents the results for a set of questions added to the midline quantitative tools. The following question was posed to head teachers:

"Now please consider a situation that might arise in your community. Imagine that two students from this community were admitted to a good university to continue their education. One student is a boy, and the

<sup>298</sup> FGD – Mothers

<sup>299</sup> KII – REOs

<sup>300</sup> KII – GWDs

<sup>301</sup> FGD – Mothers

other is a girl. But their families do not have enough money to send their children to university. Their families talk to community leaders and tried to raise money from their neighbors and friends to pay the fees. How likely do you think it is that:

- A. The family of the boy would raise enough money to send him to university?
- B. The family of the girl would raise enough money to send her to university?"

The intent behind these questions was to judge the willingness of communities to contribute money to support a boy or a girl to go to university. The graph plots the share of head teachers who thought it likely or very likely that the family of the boy/girl would raise enough money to send them. As can be seen in the graph, head teachers across the board thought boys' families were more likely to get community support for their son, but the gap between girls and boys is smaller in intervention communities, suggesting there is less bias against girls in intervention areas.

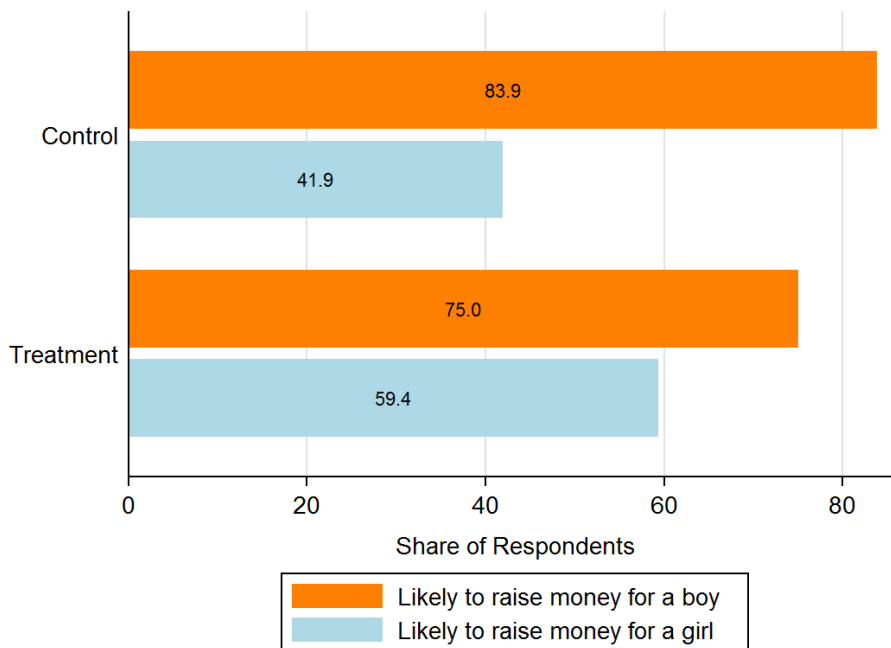


Figure 37: Head teachers' perceptions on likelihood of community support for girls' vs boys' school fees

Again, the qualitative data supports these findings. The barriers mentioned by qualitative interview participants were mentioned less frequently in the midline than in the baseline, and the most significant barriers appear to be associated with financial difficulties, as noted above. Additionally, one of the most notable findings from the story-telling exercise and other qualitative interviews is that community members are beginning to understand that girls have the option to continue their education even after they get married or have children. When presented with stories in which a girl is from an impoverished family, the first instinct of many of the girls participating in the storytelling exercises was to say that the girl might fail exams and need to drop out of school or might need to get married and get pregnant instead of continuing school. However, in some cases, girls considered other scenarios that suggest SOMGEP-T and other similar projects are having an impact on community attitudes toward girls' education. When the researcher asks whether a girl might be able to continue her education after getting married, on girl

responds, “Yes, and she managed to attend schools for mature students that are being funded by NGOs.”<sup>302</sup>

Another interesting dynamic girls appeared to consider was the level of intelligence and motivation of the girl in the hypothetical story. Girls were more likely to think that girls who were both struggling in school and from an impoverished family would drop out, with one girl reporting that such a girl would feel hopeless because her family is poor and another reporting that she would fail out of school. In stories where the hypothetical girl is praised as intelligent by her teacher or in which the girl is highly motivated to continue her education, the girls participating in the exercise appeared to feel more confident that the girl could continue her education. This was even the case when the hypothetical girl had significant responsibilities at home – many girls participating in the exercise felt confident that the hypothetical girl’s siblings or friends could help her keep up in school when she needed to miss classes, or that the teacher would be able to help raise funds for her to continue her education. Some even felt that such a girl could juggle work with schooling – attending classes in the morning and working in the afternoon. These results suggest that girls are receiving and internalizing messages about the importance of staying motivated in school. However, this finding is also notable in that it could suggest that messaging around girls’ education is encouraging the idea that only girls who are performing well should both remaining in school.

Although girls with disabilities are also internalizing these messages, with many reporting that they would like to continue their education up to the university level, evidence from their interviews and interviews with other groups suggest that disabled children face significant stigmatization. This finding will be discussed more in-depth in the section on school-related gender-based violence, but it is important to note here that girls with disability are at times verbally and physically abused in their communities and at school. Additionally, communities and schools do not appear to have the resources needed to properly accommodate children with disabilities. As one teacher explains, “...our school has not taken action to help children with disabilities because we do not know how to start such a program. In addition, we do not have the required equipment or anyone who has technical expertise on children with disabilities.”<sup>303</sup> This is supported by the findings from the interviews with girls with disabilities, as many report that they do not have any assistance getting to school and that they feel pain on their way to school and in class. For example, one of the girls who has a leg impairment reports that walking is sometimes very difficult for her, which causes her to miss school. A number of girls with visual impairments reported that their eyes are itchy in the heat, and one reports “Sometimes when I’m writing my lessons on the blackboard the lighting makes is difficult for me to see very well.”<sup>304</sup> However, despite these difficulties, the majority of the girls with disabilities interviewed in the midline reported that they are supported by their teachers, fellow students, parents, and community members. As one girl reports, “In my family, my mother supports me in my education. My teacher and classmates copy lessons for me when I’m sick.”<sup>305</sup>

## 6.6. School-related gender-based violence

A new risk mapping tool was deployed in the midline to understand where girls feel safe and unsafe in their communities, as well as the degree to which they feel safe or unsafe in these areas. Participants were given a large piece of paper and asked to work collectively to draw a map of their village, including important landmarks such as the school, shops, ALP class, and the mosque. Each participant marked her house on the map and drew the route she takes from her house to school/ALP classes. On a separate sheet of paper, girls were asked to draw where their classes are, including areas such as the teacher’s office, the toilets, and classrooms. On both of these maps – the village map and the school map – the girls were asked to mark the areas where they feel happy and secure, as well as the areas where they

---

<sup>302</sup> Story-telling – Girls

<sup>303</sup> FGD – Teachers

<sup>304</sup> KII – GWDs

<sup>305</sup> KII – GWDs

feel unsafe, scared, or at risk. The researcher then asked the girls to explain the reasoning behind their markings. For the areas where the girls indicated they felt at risk, the researcher probed to understand whether these areas are less safe for girls than boys or equally unsafe for girls and boys and why. The researcher also asked the girls what could be done to make them feel safer in these areas.

As was noted in the baseline report, sexual and gender-based violence are difficult to measure due to the sensitivity of these topics and the stigmatization associated with them. At the time of the baseline, the tools did not include questions aimed specifically at eliciting information on these topics. As a result, the topic of violence at schools was mentioned sparsely by participants, and even when it was mentioned, the topic was not addressed in detail. The only firm conclusions that could be drawn from the baseline data were the following: (1) teachers sometimes harass girls by yelling at them or using corporal punishment, (2) students sometimes fight amongst themselves, (3) people used to believe that girls would become bad as a result of mingling with boys at schools or that they would be mistreated if they were sent to school, and (4) there is concern over girls' safety on the way to school or when they are alone at home.

The new risk mapping tool was deployed in an effort to more systematically elicit information on girls' safety, as reported by girls themselves. In the aggregate, girls most often reported feeling safe and happy at school, where they can focus on studying and learning and can interact with their teachers and classmates. Girls also commonly reported feeling safe and happy in their homes, as they are free to read, do homework, and spend time with their families when they are home. Other areas mentioned less frequently by girls include the mosque, the Sheikh's house, the Madrasa (Quranic school), the road to school (where they can interact with friends), a local hotel (where they can talk to visitors), and the sea (which girls from one village visit on their days off).

By far the most commonly mentioned area where girls reported feeling unsafe is on roads/pathways to school or the market. Although girls from different areas reported different threats associated with these routes, girls across locations mentioned feeling particularly unsafe at night, with one girl reporting that she "cannot see if somebody wants to hurt me or attack me, and also no one can hear me if something happens to me because it is too far away."<sup>306</sup> In another location, a girl explains that, "There are violent things happening on the road home like rape, as well as thieves, gangs, and crazy people there at night, which increases the risk."<sup>307</sup> Girls explicitly mentioned being afraid of being kidnapped or killed in dark areas, including the mosque and the market once night falls.

There were also a number of areas mentioned by fewer respondents that are still notable. A few girls mentioned being afraid of the principal or teacher's office either because they are scared of the teachers because the office is where students get punished or because students are not allowed to visit the office without permission, or they will get in trouble. In just one of the locations, girls reported feeling unsafe because people discard vaccination needles on the road. In another location, girls reported feeling unsafe because roads are busy and some cars do not appear have any regard for the safety of students on their way to and from school. Hyenas appear to be a danger in at least one community, as they "get close at night and sometimes they eat people."<sup>308</sup>

One of the most notable findings from the midline qualitative data is that harassment by boys appears to be extremely common. The harassment takes a few different forms, one of which is verbal abuse. In one location, girls feel unsafe in the kitchen at their school because, as one girl explains, "there are boys that cause problems with us. They harass girls."<sup>309</sup> When probed, the girl added that "they say inappropriate words that I cannot tell you." The second form of harassment is physical abuse. In addition to feeling

---

<sup>306</sup> Risk-mapping – Girls

<sup>307</sup> Risk-mapping – Girls

<sup>308</sup> Risk-mapping – Girls

<sup>309</sup> Risk-mapping – Girls

unsafe on pathways at night, girls feel unsafe walking around during the day because, as one girl reports, “there are boys or men that are standing there doing bad things, so they may abuse you or throw stones at you.”<sup>310</sup> In another location, one girl reports that she feels unsafe near a certain tree that is by the Sheikh’s house because “there are boys that insult and throw stones at the girls who are walking near there.”<sup>311</sup> Lastly, girls in some areas reported that they feel unsafe at the toilets at their schools because the doors do not have locks and “there are boys who will push the door open.”<sup>312</sup> This was reported by girls across multiple schools. Girls sometimes report this type of harassment, and teachers reportedly “beat and punish” the boys, but the boys then insult the girls for telling on them rather than curbing their harassment.

Moving beyond the risk mapping exercise, there is also evidence that girls with disabilities are subjected to significant harassment and physical violence from other students. In fact, it appears that a number of the girls with physical disabilities are disabled because they were attacked. When asked about her disability, one girl explains, “A boy damaged my eye.”<sup>313</sup> Another girl explains, “A crazy man injured one of my hands so I cannot write.”<sup>314</sup> A teacher provides further evidence that disabled children are physically abused by others: “There was a blind student who used to attend school but her mother told her not to come anymore, fearing the children would further damage her eyes.”<sup>315</sup> In addition to being physically abused, girls with disabilities are verbally insulted by other students. These girls often take issues into their own hands when dealing with this abuse. As one girl explains that she is teased “sometimes, but not that much.”<sup>316</sup> She further explains, “When we are in the school some of the students call me ‘the girl who only has one eye.’ Then I attack them and they stop insulting me. The school separates us, but they do nothing else.”<sup>317</sup> One parent notes the effects this can have on disabled girls’ motivation to attend school: “Yes, in my area there is a girl who has a problem in one of her eyes and the children call her the one-eyed girl. This situation might prevent her from going to school.”<sup>318</sup> It should be noted, however, that most of the girls we interviewed reported that they have friends and feel supported at school, although it is unclear whether there was social desirability bias at play.

---

<sup>310</sup> Risk-mapping – Girls

<sup>311</sup> Risk-mapping – Girls

<sup>312</sup> Risk-mapping – Girls

<sup>313</sup> KII – GWDs

<sup>314</sup> KII – GWDs

<sup>315</sup> FGD – Teachers

<sup>316</sup> KII – GWDs

<sup>317</sup> KII – GWDs

<sup>318</sup> KII – GWDs

## 7. ALP Girls Baseline

The analysis provided in the previous two sections, especially the analysis of learning, transition rates, and the project's intermediate outcomes, has focused exclusively on cohort girls and “regular” primary schools in which SOMGEP-T is being implemented. However, in addition to interventions at primary schools, SOMGEP-T is also working to establish and support Alternative Learning Programs (ALPs) that cater to girls who are not enrolled in school. As noted briefly in the methodological description in Section 2, this evaluation serves as both a midline assessment of progress, vis-à-vis the baseline, for primary schools and cohort girls, who were selected into the evaluation's panel to be tracked from baseline through endline, and as a baseline for girls enrolled in ALPs (“ALP girls”). In future evaluation waves, ALP girls will be reassessed to study the effect of participation in ALPs on their learning scores and life skills, among other outcomes.

Data collection among ALP girls was conducted alongside fieldwork at primary schools. CARE provided a list of girls to be interviewed at ALP sample sites, based on ALP enrolment lists. In practice, Forcier researchers were unable to locate some of the girls on the sample list, and often replaced girls who could not be located with girls who were confirmed as attending the ALP in question. In total, interviews were conducted with 365 ALP girls. Interviewees completed the household survey and learning assessments in the same manner as cohort girls, with the exception that they and their caregivers were not asked questions focused on teaching quality and school management.<sup>319</sup>

In the sections that follow, we provide an overview of ALP girls: what are ALP girls like, and what sets them apart from other girls in their communities? We profile them in terms of their household and personal characteristics, as well as their learning outcomes and skills, comparing them to cohort girls and – where appropriate – cohort girls with specifically similar educational backgrounds.<sup>320</sup> This analysis serves two purposes. First, it provides insight into the type of girl who is most likely to enrol in an ALP, versus girls who drop out entirely or stay in school, which can help guide the project in targeting girls for recruitment to ALPs or other interventions. Second, it serves as a baseline against which progress for ALP girls can be measured in future evaluation waves.

### 7.1. Profile of ALP Girls

We expect girls enrolled in ALPs to be distinct from cohort girls in a variety of ways. Most fundamentally, they are not enrolled in school, while the majority of cohort girls were in school at the baseline. We expect ALP girls to be older, on average. We also expect ALP girls to be disadvantaged relative to in-school girls, facing greater barriers to educational attainment in the form of life circumstances (early marriage, born into poorer households), but it is less clear where ALP girls fit compared to OOS girls.

---

<sup>319</sup> Beyond the household survey and learning assessments, no other data collection tools – such as the head teacher survey, classroom observations or headcounts – were completed at ALPs, because ALPs are organized in a way that is fundamentally different from typical schools (i.e. there is no head teacher, per se, and classrooms are not structured as formally as at most schools).

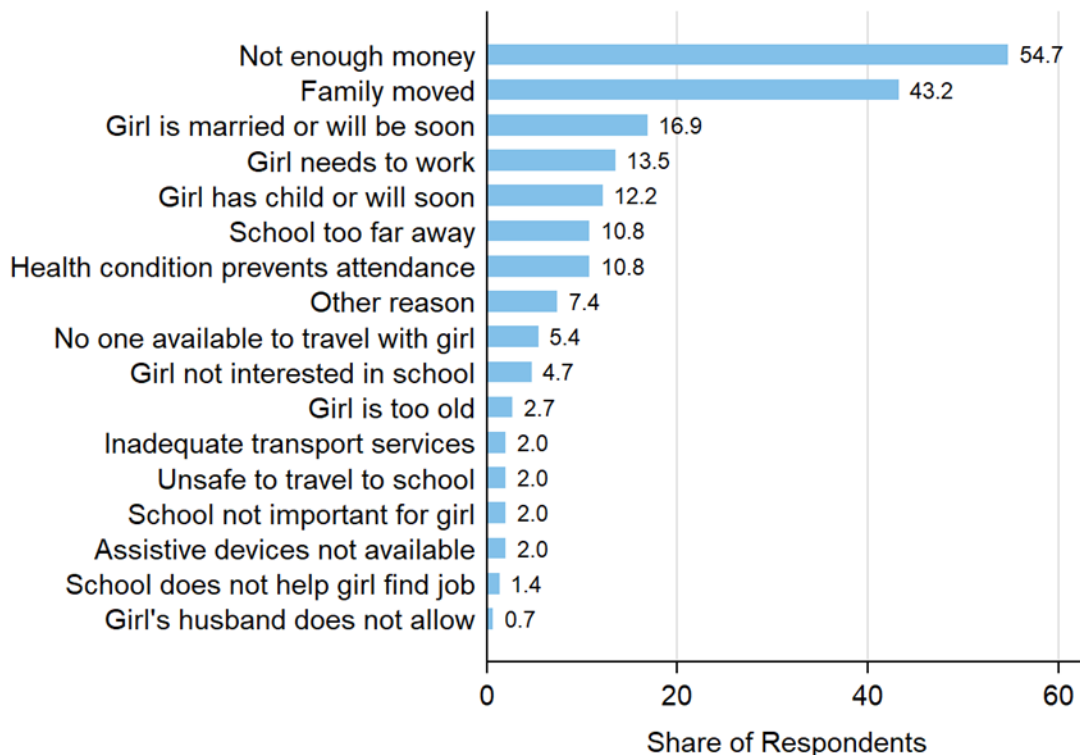
<sup>320</sup> Note that we do not assess ALP girls for transition outcomes, because they have – by definition – successfully transitioned over the past year if they are currently enrolled in an ALP school. The rate at which ALP girls drop out from the program can be assessed in future evaluation waves, if desired. We also do not assess ALPs or ALP girls on most intermediate outcomes, because metrics of school management and teaching quality were not collected for this sample.



To facilitate comparisons between both in-school and OOS girls, we use data on cohort girls from both the baseline and midline. Specifically, we compare ALP girls to in-school girls sampled at the midline, and we compare ALP girls to OOS girls sampled at the baseline (because OOS girls were not included in the midline). Our preference is for comparisons within the same time period – i.e. midline ALP to midline in-school cohort girls – to ensure that comparisons are fair. However, we are interested in understanding the difference between ALP girls and OOS girls who have not enrolled in an ALP, and use the baseline data on OOS girls to allow this analysis. Note that, when we discuss indicators that could shift significantly over a one-year time period, such as measures of economic deprivation, we encourage caution in drawing firm conclusions from the comparison of ALP girls at midline to their OOS counterparts at baseline.

Before turning to direct comparisons, it is useful to understand ALP girls’ educational backgrounds. Unfortunately, no data was collected on whether ALP girls had previously been enrolled in school or what grade level they completed before dropping out. However, ALP girls’ caregivers were asked why girls were not enrolled in school, which provides initial insight into the barriers that ALP girls face. The results from this question, which allowed respondents to select multiple reasons, are shown in the figure below. As the graph shows, the most common reasons, by far, for ALP girls to be out of school are a lack of money and migration by the family. Girls’ contemporaneous circumstances also play a role, as many respondents indicated that the girl was married or would be soon, was a mother or would be soon, or needed to work to help support herself or the family. Health considerations also played a role, as 10.8 percent of respondents cited health conditions broadly and a further 2.0 percent indicated that the girl would need assistive devices to attend school but they are not available.

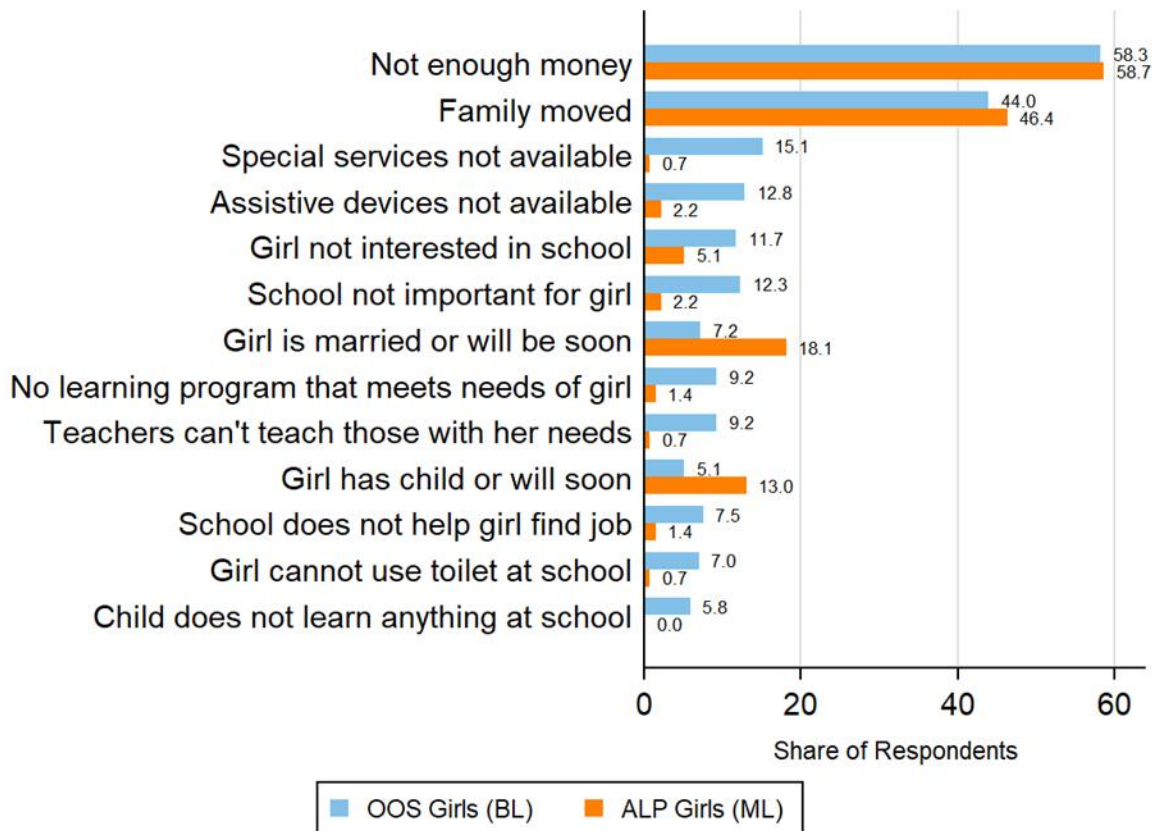
*Figure 38: Reasons girl is not enrolled in school, according to caregivers*





In terms of the most common reasons for being out-of-school, ALP girls are similar to other OOS girls in sampled communities. As shown in the figure below, a nearly identical share of OOS girl and ALP girl caregivers cited lack of money as one reason why their girls were not enrolled; a similar share noted the importance of recent household migration. But the broader sample of OOS girls diverged from ALP girls in important ways: many more OOS girl caregivers emphasized a lack of assistive devices, special services, or learning programs that their girls need. And more OOS girl caregivers indicated that they do not find education useful for their girls, reporting that schooling will not help their girl get a job or that they do not learn anything at school. In contrast, ALP girls were much more likely to be out-of-school due to marriage or childbearing – a finding that is consistent with the fact that ALP girls are older than other OOS girls, on average.

Figure 39: Reasons girl is not enrolled in school, according to caregivers, for ALP and OOS girls



How do ALP girls compare to other girls in their community more generally? The table below reports results focused on household characteristics, especially markers of relative disadvantage, such as growing up in a female-headed household, the educational attainment of one’s head of household or caregiver, and orphan status. For each indicator, we report the share of ALP girls (column 2), in-school cohort girls (column 3) and out-of-school girls (column 4) that have the characteristic in question. We also report the results of t-tests comparing ALP girls to both in-school cohort girls (column 5) and out-of-school girls (column 6), and regression models that control for the differences in age between the two samples (ALP girls versus in-school cohort girls – column 7; ALP girls versus OOS girls – column 8).

Table 79: Household characteristics of ALP, in-school and OOS girls

Indicator	Group Mean			Difference-in-Means (T-test)		Differences, Adjusted for Age	
	ALP	In-School Girls	OOS Girls	ALP vs. ISG	ALP vs. OOS	ALP vs. ISG	ALP vs. OOS
Household Size	5.3	5.4		-0.1	N/A	-0.2	N/A
Female HoH	41.4	45.2	47.1	-3.8	-5.7*	-7.3	-7.5
HoH has no occupation	41.1	47.9	48.3	-6.8*	-7.2*	-6.2	-5.2
HoH is pastoralist	11.2	7.0	13.7	4.2*	-2.4	4.4	-0.2
HoH has no formal education	63.0	69.6	80.6	-6.6*	-17.6*	-8.7	-16.1*
Caregiver has no formal education	74.0	82.8	87.8	-8.8*	-13.8*	-12.1*	-11.5*
Caregiver is illiterate	71.5	73.1	80.4	-1.6	-8.9*	-0.3	-6.5
Primary school 15+ minutes away	33.7	31.0	25.4	2.7	8.3*	5.3	14.4*
Secondary school 30+ minutes away	60.3	40.6	48.3	19.7*	12*	17.6*	2.3
Mother not in HH	21.1	13.6	19.6	7.5*	1.5	3.5	-2.1
Orphan (single or double)	14.5	13.3	11.7	1.2	2.8	1.3	0.9
Has disability	7.9	12.7	6.8	-4.7*	1.2	-3.7	6.2*

The results of this analysis run counter, in many ways, to our expectations of ALP girls. Compared to in-school girls, ALP girls are not uniquely disadvantaged by these standards: they are less likely to come from households in which the head of household has no occupation and in which the head of household has completed no formal education.<sup>321</sup> ALP girls' caregivers are more likely to have completed some form of formal education and are somewhat more likely to be literate.

The same findings broadly hold when comparing ALP girls to OOS girls. On average, OOS girls are disadvantaged relative to both in-school girls and ALP girls – their heads of household and caregivers are much less likely to have received formal education, for instance.

What distinguishes ALP girls in this realm is their relative distance from school and the presence of their mothers in their households. In total, 33.7 percent of ALP girls live 15 minutes or more from the closest primary school and 60.3 percent live 30 minutes or more from the closest secondary school. Both figures are higher than either in-school or OOS girls; in particular, ALP girls are much more likely to live a long distance from the closest secondary school. And, while ALP girls are no more likely to be orphans than their in-school counterparts, they are less likely to live in a household shared by their mother, which could reduce the willingness of their caregivers to provide for their education in some circumstances.

In line with the findings above, ALP girls do not appear to come from especially poor households, relative to in-school and OOS girls. The table below assesses each group of girls in terms of seven indicators of household wealth, capturing a range of outcomes, from durable investments (land ownership and roofing

<sup>321</sup> We define the latter group as those heads of household who have not attended school at all or have only participated in religious (madrassa) education.

material), to livestock ownership, to potentially shorter-term deprivation in the form of hunger or a scarcity of clean water.

*Table 80: Economic status of ALP, in-school, and OOS girls' households*

Indicator	Group Mean			Difference-in-Means (T-test)		Differences, Adjusted for Age	
	ALP	In-School Girls	OOS Girls	ALP vs. ISG	ALP vs. OOS	ALP vs. ISG	ALP vs. OOS
Poor Quality Roof	35.1	29.8	37.3	5.3*	-2.2	6.8	-4.7
Own large livestock	14.8	10.6	13.4	4.2*	1.4	2.4	-0.4
Own medium livestock	69.0	67.9	59.3	1.2	9.8*	3.0	8.6*
Own mobile phone	96.7	87.2	92.3	9.5*	4.4*	12.6*	4.3*
Solely own land	66.9	72.0	65.2	-5.0*	1.8	-8.0	-1.9
Went to bed hungry 1+ nights in last year	32.2	28.4	44.6	3.8	-12.3*	9.1*	-11.4*
Lacked clean water for HH use 1+ days last year	72.3	68.3	72.9	4.1	-0.6	5.3	1.2

ALP and OOS girls' households both lag in-school girls in terms of durable assets – they are less likely to own land and less likely to live in a home with a higher-quality roof. On these metrics, ALP and OOS girls are virtually indistinguishable from one another. In other ways, though, ALP girls appear to be comparatively better off than OOS girls. For example, 44.6 percent of OOS girls' households went to bed hungry at least once during the previous year, compared to 32.2 percent of ALP girls' households.

Viewed across all seven economic indicators in the table, ALP girls are never worse off than OOS girls, and – on at least three of the indicators – are ALP girls are considerably better off than OOS girls.<sup>322</sup> In contrast, the gap between ALP girls and in-school girls is less clear: ALP girls' households are less likely to have a high-quality roof, less likely to own land, and more likely to have gone to bed hungry, but those same households are more likely to own large and medium livestock and a mobile phone. It is possible that the types of economic activities in which ALP girls' households are engaged may shape these findings – the heads of ALP girls' households are somewhat more likely to be engaged in pastoralism as their primary occupation, and other differences not identified here may explain the mixed pattern of findings when comparing ALP and in-school girls.<sup>323</sup> In a very broad sense, ALP girls occupy a kind of middle ground: they are clearly better off than OOS girls but arguably worse off than in-school girls, at least on most measures.

A girls' household environment encompasses more than the demographic characteristics and economic situation of her family. Households and caregivers vary in the extent to which they prioritize education for girls, how much autonomy they provide, and the extent to which they burden girls with household chores. In the table below, we report results that focus on these more latent household characteristics. While the results elude any simple classification of ALP girls as either better or worse off than their in-school or

<sup>322</sup> Specifically, ALP girls' households are *statistically* more likely to own medium livestock, and a mobile, and less likely to have gone to bed hungry within the last 12 months. They are also more likely to own land, large livestock, and a good-quality roof, although these differences are all marginal. Important here is the consistency of the results across multiple indicators – ALP girls are better off in all cases, and substantially better off in at least a few.

<sup>323</sup> At midline, 11.2 percent of household heads in ALP girls' households reported that pastoralism was their primary current occupation, compared to 7.0 percent in households with in-school girls ( $p = 0.02$ ).

OOS counterparts, two patterns do emerge. First, in terms of caregiver *opinions*, ALP girls fare better than OOS girls and marginally worse than in-school girls. For instance, 60.5 percent of ALP girls' caregivers state that they would like their girl to complete university; this share falls below that of in-school girls (77.8 percent of whose caregivers hope they complete university) but ahead of OOS girls (51.3 percent). The caregivers of OOS girls are also the most likely to undervalue education for girls, claiming that it is not a worthwhile investment.

*Table 81: Views of girls' education and girls' decision-making power*

Indicator	Group Mean			Difference-in-Means (T-test)		Differences, Adjusted for Age	
	ALP	In-School Girls	OOS Girls	ALP vs. ISG	ALP vs. OOS	ALP vs. ISG	ALP vs. OOS
High chore burden (whole day)	28.7	9.2	51.2	19.5*	-22.5*	13.5*	-33.5*
Caregiver aspires for girl to go to university	60.5	77.8	51.3	-17.2*	9.2*	-11.4*	15.7*
Adults make schooling decisions (adult self-report)	47.1	72.5	75.2	-25.3*	-28.1*	-18.6*	-16.0*
CG says girls' education is not worthwhile	14.8	13.1	30.8	1.7	-16.0*	1.1	-16.8*
Girl has sole/joint input into schooling decisions	83.8	69.8	72.0	14*	11.8*	-0.8	1.5
Girl has sole/joint input into marital decisions	57.8	32.9	33.7	24.9*	24.1*	10.5*	13.3*

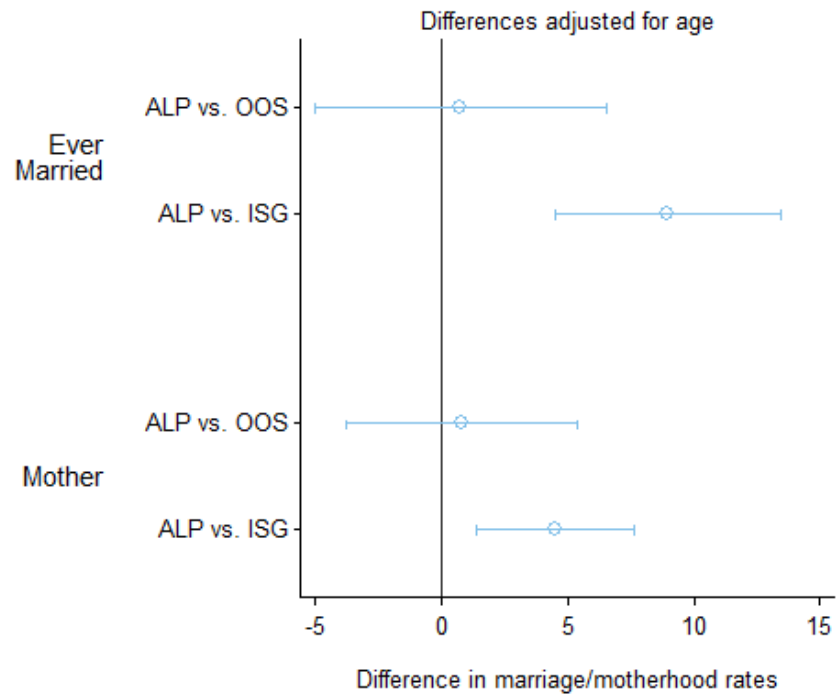
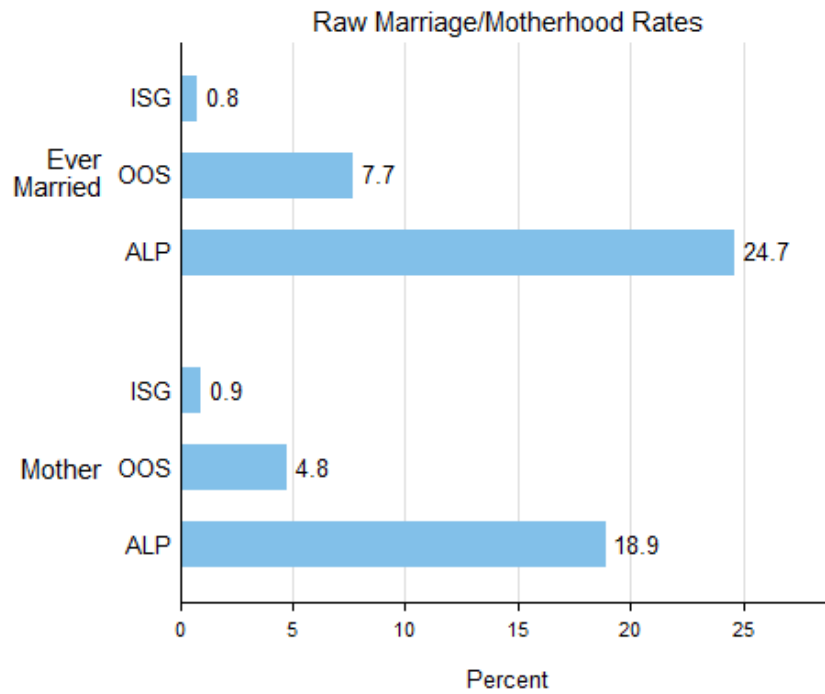
The second pattern occurs in girls' empowerment or control over important decisions. ALP girls are more likely than either of the other two groups to believe they have sole or joint decision-making power over schooling and marital decisions. ALP girls are also the least likely to live in households in which caregivers openly state that they do not take into account the opinions of their girls when making schooling decisions. To some extent, outsized influence over decisions may simply reflect the older average age of ALP girls; however, even when we control for age, ALP girls are much more likely to believe they have influence over marital decisions and their caregivers are less likely to ignore their opinions regarding schooling.<sup>324</sup> Of course, it is important to note that the relationship between enrolment in ALPs and autonomy may not be unidirectional: ALP girls may come from households that are similar to others in terms of girls' autonomy, but may feel empowered due to their participation in ALP programming.

Unlike household characteristics, ALP girls are clearly differentiated from in-school and OOS girls by their own life circumstances. Just under one-fifth (19.2 percent) of ALP girls are currently married, and nearly one-quarter (24.7 percent) are either currently married or have been married in the past. These population shares sharply diverge from in-school girls, 0.8 percent of whom are currently, or have ever

<sup>324</sup> The inclusion of married girls does not significantly affect these findings. While married girls may be more likely to believe they have control over marriage decisions – and ALP girls are more likely to be married – the results hold even when we completely exclude all girls who have been or are married.

been, married. Even compared to OOS girls, ALP girls are much more likely to have been married – 24.7 percent versus 7.7 percent.

Figure 40: Marriage and motherhood rates among ALP, in-school, and OOS girls



As noted previously, ALP girls are much older than the in-school or OOS girls in our sample (17.0 years for ALP girls, versus 13.5 years for in-school girls at midline and 13.8 years for OOS girls at baseline); given that marriage is strongly correlated with age, we also tested for differences in marriage rates between the three groups, while controlling for the age of respondents. Although controlling for age eliminates the observed difference in marriage rates between ALP and OOS girls, ALP girls are still 9 percentage points more likely to be married or have been married than in-school girls. ALP girls are also more likely to have given birth than in-school girls – even after controlling for their greater average age – but no more likely than OOS girls to be mothers. One interpretation of this finding is that ALP girls do not hail from particularly disadvantaged households but do face significant *personal* disadvantages, i.e. marriage and childbearing, that preclude enrolment in regular schools.<sup>325</sup>

In our analysis of decision-making power above, we found that ALP girls were more likely than in-school and OOS girls to have control over important life decisions. On the related outcome of self-confidence, ALP girls also appear to perform well. The table below reports the share of girls who report that they feel nervous in routine academic and social situations, such as speaking in front of an adult, or reading in public, or answering questions in class. In general, ALP girls more closely resemble in-school girls than OOS girls: 38.6 percent of ALP girls report that they are nervous speaking in front of an adult, similar to the share, 37.4 percent, of in-school girls who feel the same. In contrast, 62.4 percent of OOS girls report feeling nervous speaking in front of an adult; a similar gap occurs when respondents were asked to rate their nervousness when speaking in front of a group of their peers. As in our discussion of empowerment and decision-making power, these results should not be construed to mean that more confident girls are more likely to enrol in ALPs – it is possible that participation in an ALP could increase self-confidence, especially in academic situations, though our data do not allow us to distinguish between these two potential explanations.

Table 82: Life skills and self-confidence of ALP, in-school, and OOS girls

Indicator	Group Mean		Difference-in-Means (T-test)		Differences, Adjusted for Age		
	ALP	In-School Girls	OOS Girls	ALP vs. ISG	ALP vs. OOS	ALP vs. ISG	ALP vs. OOS
Nervous when speaking in front of an adult	38.6	37.4	62.4	1.3	-23.8*	4.7	-20.0*
Nervous when speaking in front of peers	21.6	25.7	57.6	-4.1	-35.9*	-4.4	-34.1*
Nervous reading in front of others	36.6	35.6	N/A	1.0	N/A	4.2	N/A
Nervous doing maths in front of others	37.5	34.0	N/A	3.5	N/A	5.2	N/A
Confident answering questions in class	89.9	89.9	N/A	0.0	N/A	-3.9	N/A
Confident answering questions in group of people	86.8	86.8	82.9	0.0	4.0*	-4.0	0.3

<sup>325</sup> While the Federal Government of Somalia does not bar pregnant girls or mothers from enrolling in school, there is no law guaranteeing their access to education, nor are there explicit policies that encourage continued enrolment or re-enrolment after giving birth (see: Human Rights Watch. 2018. “Leave No Girl Behind in Africa: Discrimination in Education against Pregnant Girls and Adolescent Mothers.” Available at [https://www.hrw.org/sites/default/files/report\\_pdf/au0618\\_web.pdf](https://www.hrw.org/sites/default/files/report_pdf/au0618_web.pdf)).



## 7.2. ALP Girls' Learning Outcomes

This section presents learning outcomes for ALP girls. Because this is the first time that data has been collected from ALP girls, this section of the report functions as a baseline for this population, and the analysis is limited to reporting on baseline values (with comparison to cohort girls as a reference point) and establishing foundational learning gaps for the purpose of targeting learning-based interventions in the future.

### Literacy

The table below summarises average ALP literacy scores by age-band because ALP girls are not assigned to conventional grade-levels. In order to provide a basis for comparison, the second column in the table presents the average scores for in-school, cohort girls belonging to the same age-band as the relevant ALP girls. In the aggregate, the average literacy score for ALP girls is 46.3, which is 12.5 percentage points lower than the average literacy score for in-school, cohort girls.

Table 83: ALP Literacy (EGRA/SeGRA)

Age	ALP Mean	Cohort Mean	Difference (ALP - Cohort)	Standard Deviation for ALP
11-13	36.3	54.6	-18.2	31.1
14-15	33.4	62.4	-29.0	31.5
16-17	48.7	64.3	-15.6	31.4
18-19	52.7	74.9	-22.2	32.9
20-21	50.6	83.3	-32.7	31.2
Overall	46.3	58.8	-12.5	31.4

Taking ALP scores on their own, table shows that literacy learning generally increases as a function of age, but that there is not a monotonic increase from one age-band to the next. For example, learners from 11-13 years of age have a somewhat higher score than their older peers who are 14-15 years of age. Because all ALP girls were out of school for varying periods of time prior to entering ALPs, their learning levels are extremely varied, and older girls may have been out of school longer than younger girls (leading to differential levels of skill loss due to lack of practice and forgetting), and each girl may have dropped out at a different grade-level meaning that the literacy skills she attained prior to dropping out will be different as well.

At each age-band, ALP girls score much lower on average than in-school cohort girls. At baseline, this is to be expected, and one potential reference-point for improvement during the next midline will be if ALP girls can close the gap with cohort girls of the same age. In terms of current differences between ALP and cohort girls, the largest learning gaps in literacy are with girls aged 14-15 years (-29.0 percentage point gap), and with girls aged 20-21 years (-32.7 percentage point gap). These findings suggest that there may be specific and severe skill gaps for ALP girls in these age-bands that account for their extremely low performance vis-à-vis cohort girls of a similar age.

### Numeracy

The table below summarises average ALP numeracy scores by age-band in the same fashion as literacy scores summarised above. In the aggregate, the average numeracy score for ALP girls is 47.1, which is 4.0 percentage points lower than the average numeracy score for in-school, cohort girls.

Table 84: ALP Numeracy (EGMA/SeGMA)

Age	ALP Mean	Cohort Mean	Difference (ALP - Cohort)	Standard Deviation for ALP
11-13	36.0	46.4	-10.4	25.4
14-15	37.6	54.9	-17.3	26.5
16-17	48.2	58.2	-10.0	26.3
18-19	50.9	64.2	-13.3	25.4
20-21	54.6	78.0	-23.3	25.0
Overall	47.1	51.1	-4.0	26.0

As with literacy, the table shows that numeracy scores for ALP girls increase as a function of age. Unlike literacy scores, the numeracy scores do increase monotonically by age despite the differential dropout grades and durations of time out of school that were noted above. This finding provides further support for the hypothesis that literacy skills can be lost more quickly than numeracy skills. In this same line of reasoning, the overall gap between ALP and cohort girls is much smaller for numeracy than for literacy (-4.0 for numeracy versus -12.5 for literacy).

As above, ALP girls score lower on average than in-school cohort girls within each age-band. Despite the fact that numeracy learning appears to be less attenuated for ALP girls than their literacy learning, the same girls in the same age-bands demonstrate the largest gaps in performance when compared with in-school, cohort girls in the same age-band. As with literacy, the largest learning gaps in numeracy are with girls aged 14-15 years (-17.3 percentage point gap), and with girls aged 20-21 years (-23.3 percentage point gap). Again, these findings suggest that there may be specific and severe skill gaps for ALP girls in these age-bands that account for their extremely low performance vis-à-vis cohort girls of a similar age. These skill gaps will be explored in greater detail in the following section.

#### *Identifying Foundational Skill Gaps*

The tables on the following page present foundational skill gaps for numeracy. As expected, the percentage of non-learners generally increases as a function of increasing sub-task difficulty, while the percentage of proficient learners decreases correspondingly. At higher levels of difficulty, the distribution of achievement levels shows evidence of moderate bimodality with the number of emergent and established learners becoming somewhat smaller at higher skill levels. For advanced word problems (involving multiplication and division), the distribution is strongly bimodal.

Table 85: ALP Foundational numeracy skills gaps

Categories	Subtask 1 Number Identification	Subtask 2 Addition (Level 1)	Subtask 3 Subtraction (Level 1)	Subtask 4 Addition (Level 2)	Subtask 5 Subtraction (Level 2)	Subtask 6 Word problems (add/subtract)	Subtask 7 Multiplication (Level 1)	Subtask 8 Multiplication (Level 2)	Subtask 9 Division (Level 1)	Subtask 10 Division (Level 2)	Subtask 11 Word problems (mult/div)
<b>Non-learner 0%</b>	6.3	9.6	12.6	26.6	32.3	15.3	38.6	80.3	60.3	86.3	56.7
<b>Emergent learner 1%-40%</b>	44.9	3.0	4.1	13.7	9.3	5.8	12.9	4.7	14.2	5.2	0.0
<b>Established learner 41%-80%</b>	29.9	19.5	14.2	23.6	22.7	32.1	28.2	6.0	12.6	3.6	18.6
<b>Proficient learner 81%-100%</b>	18.9	67.9	69.0	36.2	35.6	46.8	20.3	9.0	12.9	4.9	24.7
	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

The primary learning gap in numeracy is between level-1 subtraction and level-2 addition (subtask 3 to subtask 4). At this point, the proportion of non-learners increases by 14.0 percentage points, while the proportion of proficient learners decreases by 32.9 percentage points. This skill gap indicates that, while the majority of ALP learners are proficient in arithmetic operations of addition and subtraction, they have trouble applying these operations to larger numbers with multiple digits where “carrying” is required. Roughly a quarter of ALP girls fall into the emergent or established learner categories for the harder addition and subtraction tasks, suggesting that they fundamentally have the necessary skills and will continue to improve through additional practice. Another quarter of ALP girls fall into the non-learner category. These individuals may be at risk of falling behind if they do not receive additional remedial work.

Another even more severe learning gap exists between level-1 multiplication and level-2 multiplication (subtask 7 to subtask 8), with the proportion of non-learners increasing by 41.6 percentage points, and the proportion of proficient learners decreasing by 11.2 percentage points. This is also the point at which the distribution of proficiency levels becomes somewhat bimodal. It is likely that most ALP girls, because they were out of school for at least one year, have either forgotten or simply never learned advanced multiplication (or even simple multiplication, for that matter), and thus this skill will improve naturally as they are introduced or re-introduced to multiplication skills in the context of ALP learning.

As with cohort girls, ALP girls’ performance on word problems is at a similar level to their performance on the fundamental skills necessary for solving those problems, meaning that girls who are able to solve addition and subtraction problems also tend to be equally able to solve word problems that employ those operations.

The tables on the following page present foundational skill gaps for literacy. As with numeracy, the percentage of non-learners generally increases as a function of increasing sub-task difficulty, while the percentage of proficient learners decreases correspondingly. At higher levels of difficulty, the distribution of achievement levels shows evidence of severe bimodality with the number of emergent and established learners becoming far smaller at higher skill levels.

Table 86: ALP Foundational literacy skills gaps

Categories	Somali ST1	Somali ST2	Somali ST3	Somali ST4	Somali ST5	Somali ST6	Somali ST7	Somali ST8
	Reading Words	Reading Comp (easy)	Reading Comp (medium)	Reading Fluency	Reading Comp (difficult)	Writing (fill blank)	Writing (negative form)	Writing (future tense)
<b>Non-learner 0%</b>	13.7	25.5	31.8	19.5	47.7	44.9	54.5	58.9
<b>Emergent learner 1%-40%</b>	9.6	8.2	9.0	25.2	17.3	13.2	3.3	3.8
<b>Established learner 41%-80%</b>	13.4	37.8	31.8	31.2	28.2	14.8	6.8	6.3
<b>Proficient learner 81%-100%</b>	63.3	28.5	27.4	24.1	6.8	27.1	35.3	31.0
	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

The primary skill-gap in ALP literacy almost immediate. While the majority of ALP girls being proficient in reading words (subtask 1), but only 28.5 percent being able to understand what they are reading in each reading comprehension questions (subtask 2). From subtask 1 to subtask 2, there is a 34.8 percentage point reduction in the proportion of proficient learners and an 11.8 percentage point increase in the proportion of non-learners. Within this skill gap, a plurality of learners still fall into the established learner category for easy reading comprehension, which suggests that they will continue to improve with more practice. However, the 25.5 percent of learners who fall into the non-learner category for easy reading comprehension are likely to remain stuck at this level without significant remedial work.

There is a second skill gap from medium to difficult reading comprehension (subtask 3 to subtask 5). At this point, there is a 20.5 percentage point decrease in the proportion of proficient learners and a 15.9 percentage point increase in the proportion of non-learners. It is unclear whether or not ALP girls falling into this skill gap require remedial work, but there is potential for their proficiency to improve naturally as a result of improvements in more fundamental reading comprehension skills (i.e. at the level of subtask 2).

Finally, writing skills appear to be keeping pace with reading skills. The proportion of proficient learners in writing is roughly equivalent to the proportion of proficient learners in easy and medium levels of reading comprehension. However, the proportion of non-learners is consistently higher for writing tasks as compared with reading tasks of equivalent difficulty.

#### *ALP – Grade Levels Achieved*

This section reports the share of girls achieving each grade level of numeracy, based on an analysis of the school curricula or syllabi. The section above on cohort grade-levels achieved has rehearsed the details of how achievement levels were coded at baseline. For ease of reference, the table below shows grade-level standards for numeracy.

*Table 87: Grade Level Standards for Numeracy*

<b>Grade Level Achieved</b>	<b>Numeracy Skills</b>
<b>1</b>	<ul style="list-style-type: none"> <li>• Number identification up to 99 (portion of subtask 1)</li> <li>• Addition without carrying numbers (portion of subtask 2)</li> <li>• Subtraction without borrowing (subtask 3)</li> </ul>
<b>2</b>	<ul style="list-style-type: none"> <li>• Number identification up to 999 (portion of subtask 1)</li> <li>• Addition carrying one number (portion of subtask 2)</li> <li>• Addition with 3 digits, carrying up to 1 number (subtask 4)</li> <li>• Subtraction carrying one number (portion of subtask 5)</li> <li>• Addition and subtraction word problems with simple underlying arithmetic (subtask 6)</li> <li>• Multiplication of 1-digit numbers (subtask 7)</li> <li>• Division of 2-digit number by 1-digit number (subtask 9)</li> </ul>
<b>3</b>	<ul style="list-style-type: none"> <li>• Subtraction carrying two numbers (portion of subtask 5)</li> <li>• Multiplication of 2-digit numbers (subtask 8)</li> <li>• Word problems with simple multiplication and division (subtask 11)</li> </ul>
<b>4</b>	<ul style="list-style-type: none"> <li>• Division of 3-digit number by 2-digit number (subtask 10)</li> </ul>
<b>5</b>	N/A

Because ALP girls are not assigned to conventional grade-levels, their achievement levels are analysed with reference to their age, with ALP students being grouped into the same age-bands as were employed in the foregoing analysis of ALP learning outcomes. In the table below, the achievement levels for ALP girls are provided with each cell presenting the percentage of girls within a given age-band (columns) who have achieved a given grade-level in terms of their performance on relevant subtasks (rows).

*Table 88: Grade level achieved by ALP girls in numeracy*

	11-13	14-15	16-17	18-19	20-21
Grade 1 achieved	17.4	27.0	25.0	29.6	35.0
Grade 2 achieved	0.0	2.7	4.0	1.9	6.7
Grade 3 achieved	0.0	0.0	5.0	1.9	1.7
Grade 4 achieved	4.3	2.7	2.0	6.5	3.3

The primary finding that emerges from this analysis is that the majority of ALP girls in all age-bands are learning at or below a grade-1 achievement level. These results parallel the results for in-school cohort girls, insofar as most girls are performing at an achievement level that is far below the grade-level that they would be attending (based on their age) if they were in school. It is also worth noting, among the oldest girls, that their achievement levels are somewhat bimodal – being split between those who have reached a grade-4 achievement level and a plurality who have only reached a grade-1 achievement level, with comparatively few older learners having achieved at grade-2 or grade-3 levels.

### 7.3. ALP – Subgroup analysis of the Learning Outcomes

This section presents an analysis of learning outcomes by key subgroups of the population of ALP girls. Because this is effectively a new baseline for ALP girls, the table below presents a baseline-style summary of average literacy and numeracy scores by key subgroups that were identified as salient during the SOMGEP-T baseline study of cohort girls. A few subgroup categories (e.g. hearing disability) are left out of this table and analysis because there are no ALP girls who belong to those subgroups. Ultimately, many of the subgroups are exceedingly small due to the small overall sample size of ALP girls. Nonetheless, a few important subgroups emerge as predictors of learning outcomes, and these salient subgroups tend to also be congruent with the subgroups that were strong predictors of learning outcomes among cohort girls.

	Average literacy score (aggregate)	Average numeracy score (aggregate)	Number of observations for subgroup
<b>Characteristics:</b>			
All in-school girls	46.3	47.1	365
Living without both parents	46.8	47.7	64
<b>Disability</b>			



Vision impairment	75.5*	47.3	1
Mobility impairment	3.9*	6.4*	1
Mental health impairment	54.2	48.1	28
Anxious	55.9	52.3	20
Depressed	56.9	47.2	19
Any disability	54.9	48.1	29
<b>HOH and Carer Characteristics</b>			
HOH no wage-earning occupation	43.2	45.3	165
HOH no education	43.0	42.9	153
HOH female	46.7	46.9	151
HOH Pastoralist	30.5*	36.2*	41
Carer no education	42.3	44.1	167
<b>Household Assets</b>			
Owns camels	35.0*	39.6*	55
Owns medium-sized livestock	43.9	44.4*	252
Owns small livestock	44.2	36.1	15
Owns mobile phone	46.4	47.1	353
Access to water reservoir/storage	50.0*	48.7	243
Owns land	47.0	46.8	243
<b>Poverty</b>			
House is informal/temporary structure	42.9	46.5	21
Gone to sleep hungry most days	21.6*	30.8	8
Gone without enough clean water most days	32.3*	41.3	33
Gone without medicines or medical intervention most days	36.1*	43.0	77
Gone without cash income most days	44.0	46.4	99
<b>Migration</b>			
Displaced or moved in past 12 months	39.8	39.1	4
Household migrates seasonally	39.3	42.2	12
<b>Other</b>			
High chore burden (whole day spent on chores)	50.4	50.2	103
Married	47.7	47.5	70

Mother, under 16	43.2	28.0*	2
------------------	------	-------	---

\*Note, an asterisk indicates results that are statistically significant at the 95% confidence level (or higher) in a regression with cluster-robust standard errors.

In keeping with the findings regarding cohort girls above, several proxies of pastoralist lifestyle are strong predictors of lower learning outcomes. Girls with heads of household who reported their profession as being pastoralist, as well as girls belonging to households that own camels (a proxy for an itinerant/pastoralist lifestyle) had literacy and numeracy scores that were significantly lower than average. Ownership of medium-sized livestock is also predictive of lower learning outcomes (to a statistically significant degree for numeracy). Taken together, these findings provide strong evidence that pastoralism is associated with ALP girls having lower learning outcomes, which is in keeping with the findings related to cohort girls and the attendant qualitative analysis above.

While poverty was not a strong predictor of lower learning scores among cohort girls, there are three key proxies of household poverty that are significant predictors of lower learning outcomes among ALP girls. Girls whose caretakers reported that the members of their household go to sleep hungry most days, often go without clean water and often go without needed medical attention had lower literacy and numeracy scores (with literacy scores being lower to a statistically significant degree for all three poverty proxies).

The analysis of girls with physical disabilities has sample sizes that are far too small to allow for strong conclusions to be drawn, but it is clear that the girl who identified as having a mobility impairment is performing far below average and is likely to be at high risk of falling irreversibly behind and also potentially dropping out of school.

To briefly summarise the key findings above, ALP girls are, by their very nature, at higher risk of underperformance and dropout than in-school cohort girls. At any given age, ALP girls have average learning scores that are below their peers in the cohort sample. ALP girls face many of the same challenges and barriers as cohort girls – most notably the effects of violent conflict, and the disproportionate effects of drought on pastoralist households. Finally, poverty is a major predictor of underperformance for ALP girls. The strong effects of poverty on ALP learning are worth noting because many girls probably ended up being sorted into ALP programs due to poverty. The fact that poverty remains a significant predictor of outcomes even within the ALP sample thus implies that even within the population of at-risk girls there are important variations in household economic distress that can further impede learning.

## 8. Conclusion & Recommendations

### 8.1. Conclusions

#### Learning Outcomes

For cohort girls, Somali literacy scores for the intervention group were 42.9 at the baseline and 59.8 at the midline (+16.9 percentage points from the baseline); scores for the comparison group were 40.3 at the baseline and 58.5 at the midline (+18.2 percentage points from the baseline).<sup>326</sup> Progress against the midline literacy target of 6.4 percentage points (over and above comparison) is -20 percent due to the fact that the amount of improvement in the comparison group was slightly greater than the amount of improvement in the intervention group (but not to a statistically significant degree). Numeracy scores for the intervention group were 40.0 at the baseline and 51.9 at the midline (+11.9 percentage points from the baseline); scores for the comparison group were 38.0 at the baseline and 50.6 at the midline (+12.6 percentage points from the baseline).<sup>327</sup> Progress against the midline numeracy target of 6.5 percentage points (over and above comparison) is -12 percent due to the fact that the amount of improvement in the comparison group was slightly greater than the amount of improvement in the intervention group (but not to a statistically significant degree). In addition, it was determined that the numeracy boost training delivered to teachers during the first phase of implementation had not yet had a statistically significant effect on numeracy outcomes among the girls being taught by those teachers.

When considering negative progress against midline targets for literacy and numeracy, readers should note that, this finding does *not* mean that girls in the comparison group out-performed girls in the intervention group. In an absolute sense, intervention girls still had average learning scores that were higher than the average scores for the comparison group. Girls in both intervention and comparison groups improved over time about as much as we would expect given the fact that the girls were attending school, and thus learning. This is the normal trajectory of learning that one would expect to observe in the absence of any additional intervention. It is a statistical fact that the amount of improvement over time was slightly higher for the comparison group than for the intervention group, but the DID Beta is less than 1%. That result is statistically and substantively indistinguishable from zero, which implies zero measurable intervention-effect, and nothing more.

These findings related to learning, as well as the findings related to outcomes below should all be interpreted in light of the fact that the implementation of the intervention occurred for only four months prior to the midline measurements that form the basis for this analysis. Thus, a full year intervened between the time of the baseline and the midline, but beneficiaries received only about one third of a year's worth of exposure to project interventions. The lack of measurable, positive intervention-effects thus far should not be viewed as problematic or as indicating the likelihood that the intervention will not have the desired effect once the duration of exposure has increased.

While the intervention as a whole did not have a detectable positive effect at baseline, it was found that girls who reported participating in the Girls' Empowerment Forum (GEF) had significantly higher literacy and numeracy scores than their peers who did not report participating in GEF. This finding suggests that

---

<sup>326</sup> Note, these figures include grade 3 (3) girls who were later excluded in difference-in-differences estimation and the calculation of progress against midline targets.

<sup>327</sup> Note, these figures include grade 3 (3) girls who were later excluded in difference-in-differences estimation and the calculation of progress against midline targets.

at least some project interventions such as support for GEFs are beginning to have a measurable positive effect, even if the overall effects of the intervention are not yet evident.

Girls from pastoralist households were identified as vulnerable at the baseline, and they remain particularly disadvantaged at the midline, tending to have significantly lower literacy and numeracy outcomes than their peers. Qualitative data suggests that pastoralist households were hit the hardest by the effects of recent drought, leading to increased economic distress among pastoralist households that affected girls' abilities to have adequate study time and to attend school regularly. In contrast to baseline findings, disability status was no longer a predictor of significantly lower learning outcomes, with the exception of reported vision impairment being consistently associated with lower literacy and numeracy outcomes (to a statistically significant degree for numeracy). . In terms of barriers that girls face, school and teaching quality were identified as problems at the baseline, and at midline poor principal and poor teacher performance remain two of the strongest predictors of lower learning outcomes. Despite a decreasing performance gap between boys' and girls' learning, unequal treatment of boys and girls by teachers in the classroom is strongly associated with lower learning outcomes.

### Transition Outcomes

Overall, transition rates increased markedly in intervention communities relative to their baseline levels. While we have frequently noted the shortcomings of our transition analysis, the comparison of like-for-like samples of baseline and midline girls gives us confidence that the upward shifts in transition rates at midline are not merely a consequence of sampling variation. Using the identified subsamples, we found that transition rates increased by 5.9 percentage points (from 76.6 percent to 82.5 percent) in intervention communities from baseline to midline. The 5.9 percentage point improvement in intervention communities, though focusing only on a specific *type* of cohort girl, is still a significant improvement. At the same time, comparison communities have experienced a similar improvement in transition, which complicates statements about causality or attribution of the change to the project or its activities.

Among the subpopulation studied, transition pathways were limited – very few girls enrolled in alternative education, vocational training, or entered employment. But even within the limited pathways that were common, trends emerged: repeating a grade was much more common among younger girls; among older girls, it appears that those who might otherwise repeat a grade due to falling behind instead drop out of school.

Disaggregated analysis of transition showed large shifts in transition rates among subgroups from baseline to midline. Many of the girls who were most disadvantaged at baseline – those with mental health disabilities, and facing a significant chore burden at home – still lag behind their peers, but have seen large improvements. Transition rates among girls with mental health disabilities improved by 18.8 points in intervention communities, for instance, and girls who are responsible for a full day of household work saw a 23.5 point increase in their transition rate vis-à-vis the baseline (to 71.9 percent among intervention communities).

### Intermediate Outcomes findings

#### *Attendance*

The attendance rates of girls in the intervention have not improved after one year of intervention. Girls in intervention areas had a headcount attendance rate of 77.0 percent at midline, a slight decrease from the 82.2 percent headcount attendance rate of intervention girls in the baseline. This decrease in intervention areas however was not as large as the decrease in the headcount attendance rates of girls in comparison areas which fell from 83.9 percent to 78.0 percent in the midline. As such, difference in difference

analysis found a positive average intervention effect for the intervention over time, but this effect was not significant.

Once girls and boys are enrolled in school, there is not a gap in attendance between girls and boys. This seems to suggest that current interventions with community members, teachers, and parents are sufficient in building support for sending girls to school. As was found in the baseline, the study did not observe a major attendance gap between girls and boys at the midline. Boys had a headcount attendance rate of 82. percent in the baseline and 79.3 percent in the midline.

The school survey attendance rates for intervention areas in the midline is 91.7 percent, a marginal decrease from the baseline 91.9 percent attendance rate recorded in the baseline. Comparison schools experienced a similar drop in the school survey attendance rates, from 94.5 percent in the baseline to 94.3 percent in the midline. A significant positive intervention effect is not observed in difference in difference modelling.

Better data is being gathered about girls' attendance in schools. However, there was a significant improvement from baseline to midline in attendance recordkeeping that was captured by the school survey. Only 49.1 percent of cohort girls had enrolment records in the baseline while 70.3 percent of cohort girls had enrolment records in the midline. The intervention effect is positive, but not significant. In addition to more enrolment records being kept in the midline, it seems as if the accuracy of the recordkeeping has also improved as the relationship between the school survey and classroom headcounts has improved from baseline to midline.

#### *School Governance and Management*

From the head teachers' assessment, it was found that CEC activities (such as monitoring teachers' attendance, facilities, teaching quality, or students' attendance and drop out during their last visit) in the intervention group did not significantly differ from the comparison group as a result of participating in the SOMGEP-T project. The existence of school/CEC management plans has increased from 28.5 percent to 46.0 percent overall, but when data is disaggregated by intervention, it is evident that most of this increase has taken place among the comparison group. The percentage of head teachers in the intervention group reporting having a CEC plan has increased from 43.7 percent in the baseline to 53.1 percent in the midline, while this has tripled among CECs in the comparison group from 12.9 percent to 38.7 percent.

Yet, this differences between the intervention and the comparison group is not statistically significant. Also, the project did not seem to have significant impact on the total amount that CECs contribute to teachers' salaries nor on whether teachers received incentives to stay in school.

From the caregivers' perspectives, the CEC appear to have better performance, and monitoring student attendance happens to be the most frequent CEC action cited by the parents (64.3 percent), followed by monitoring teacher attendance (51.0 percent), and improving school infrastructure (22.6 percent). The only significant difference-in-differences that was observed between the intervention and the comparison group was the provision of financial support to students where CECs in the intervention group have increased their financial support to students from 9.5 percent in the baseline to 16 percent in the midline, while it has dropped from 10.4 percent to 4.1 percent among the schools in the comparison group.

#### *Teaching Quality*

Across a wide range of indicators, teaching quality has improved from baseline to midline in both intervention and comparison schools. In intervention communities, 85.7 percent of caregivers believe that

their girl's teacher(s) have improved over the past year. Teachers in intervention schools are now more likely to ask open-ended questions, specifically reach out to a student who has not been participating, and provide encouraging feedback. They are simultaneously less likely to use rote copying from the board as a teaching method and less likely to use corporal punishment.

Improvements in teaching quality are, in many cases, shared between intervention and comparison communities. After accounting for the gains made in comparison communities, the project's impact on teaching quality is mixed and fairly limited – a small net improvement in classroom participation, but a *relative* increase in the use of corporal punishment during classroom observations, for instance. Even findings surrounding this key individual indicator – use of corporal punishment – is complicated by contradictory results: intervention schools logged a relative increase in observed use of corporal punishment during classroom observations, but a relative decrease in student reports of corporal punishment by their teachers. Drawing conclusions in some cases requires adjudicating between two imperfect measures of the outcome. Where the project does seem to have had a sizable impact on intervention schools, specifically, is in an area that teacher training can impact most readily: the use of formative assessments. The share of teachers in intervention schools using formative assessments rose from 43.1 to 71.9 percent in the last year, compared to a more modest 14 point increase in comparison schools.

### *Life Skills*

It was found that girls' leadership and life skills have slightly improved as a result of their participation in the SOMGEP-T project. Although girls in the comparison group gained a higher average YLI score compared to that of the girls in the intervention group, the girls in the intervention group have (excluding the younger out-of-school girls) more confidence in their life skills by scoring higher than the comparison group. Particularly, girls in the intervention group generally have more confidence speaking in front of an adult or in a group of people as well as doing math and answering questions in class. Yet, none of these differences were statistically significant. The only subgroup girls who had significant difference in difference in their level of confidence speaking in front of a group was the girls with disabilities.

The qualitative evidence collected from the FGDs with mothers and teachers and KIIs with girls with disabilities also suggests that girls' participation in decision-making and engagement in their homes, schools, and communities is improving, but there are some barriers that girls face. There is also a lack of consensus among mothers regarding whether their girls feel comfortable expressing their ideas at home, school and in the communities. Girls' shyness at home, in school, and in the communities was also a theme that emerged across groups. From the girls' perspective, they feel confident speaking in their homes, school, and communities despite the challenges that girls with disabilities face.

### *Community-based attitudes and behaviour change*

Community attitudes toward girls' education appear to have improved since the baseline, with both quantitative and qualitative evidence suggesting that these improvements are as a result of the project. For quantitative indicators of community attitudes, improvements since the baseline have been more substantial in intervention areas than in comparison areas. In the qualitative interviews (which were conducted only in intervention areas), there is almost complete consensus among participants on the value of girls' education, representing a positive shift since the baseline, during which participants were actively debating the value of girls' education relative to the value of boys' education. Although the same barriers to girls' education were mentioned in the midline that had been mentioned in the baseline (e.g. early marriage, prioritization of boys' education over girls' education), there is evidence that girls are internalizing messaging on the importance of remaining motivated in school, and the responses of a number of participants suggest that girls and community members understand that there are options for

girls to continue their education even after they get married and have children. Girls with disabilities also appear to be internalizing these messages, with many aspiring to continue their education up to the university level if possible. However, it should be noted that the storytelling exercise with girls revealed that girls are more likely to think girls who are from an impoverished family and struggling in school will drop out than girls who are from an impoverished family but are intelligent or motivated in school. This finding may suggest that messaging around girls' education is being misinterpreted in some cases and encouraging the idea that only girls who are performing well should bother staying in school.

Data gathered in this evaluation round suggests that there needs to be an increased focus on girls with disabilities, particularly around facilitating their access to material support, establishing child protection policies that are sensitive to the unique needs of children with disabilities, and dispelling harmful community attitudes toward children with disabilities. In the qualitative interviews, girls with disabilities often reported that they have difficulty reaching school, regularly attending school, and focusing in school as a result of their disabilities. For example, girls who have trouble walking may not have access to wheelchairs or crutches, and girls with eye impairments do not all have access to treatment and complained that their eyes are itchy in the heat of the sun, that they suffer from headaches, and that they have trouble seeing well in class. The qualitative interviews with CEC members, mothers, and teachers suggest that not much is being done in schools or communities to assist disabled girls – fundraising tends to focus more on covering school fees for marginalized children, and respondents explained that they do not have the technical expertise to help girls with disabilities. The findings suggest that the challenges girls with disabilities are facing affect not just their access to education, but also their sense of inclusion in their communities and in school. Girls with disabilities are often the subject of verbal, physical, and emotional abuse from other students and community members, and there do not appear to be any child protection policies in place that would ensure girls with disabilities feel safe at school. In fact, girls with disabilities reported in some cases having become disabled as a result of physical abuse or getting into fights to defend themselves from abuse they face because of the stigma associated with having a disability.

#### *School-related gender-based violence*

In the midline, the new risk mapping exercise allowed for a more in-depth analysis of issues surrounding girls' safety than was possible in the baseline. In this exercise, girls most commonly reported feeling happy and safe at school and in their homes and feeling unsafe or at risk on roads and pathways to important village points, like the school or market. Girls feel particularly unsafe in these areas at night, when thieves, gang members, or others who are on the road might harm them. Although girls most often reported feeling safe at school, there are some areas in school, and in the community, where girls do not feel safe because they are harassed by boys. Girls reported boys harassing them by throwing stones at them, saying inappropriate things to them, and opening the toilet doors. In the mapping area, girls mentioned a number of other areas in which they feel safe or unsafe, but these areas were specific to their villages. This is an important finding in itself, in that it suggests that there may be a need for individual risk mapping exercises in each target school in order to properly tailor approaches to curbing violence or insecurity more generally.

As was noted above, girls with disabilities appear to face even more significant levels of harassment from other students, and often feel obligated to take matters into their own hands when teachers are unhelpful in curbing teasing. In some cases, girls with disabilities reported that they became disabled as a result of being attacked, and there was additional evidence that girls with disabilities might be avoiding school for fear of being attacked.

#### **Project approach to gender inequality**



The project continues to take a robust approach to the measurement of gender inequalities and the understanding of vulnerable subgroups in terms of learning outcomes as well as intermediate outcomes. The midline findings reveal that there is still an aggregate-level gap between girls' and boys' learning outcomes, with boys having higher average scores than girls in both literacy and numeracy. However, girls appear to be closing the gap and girls in grade 3 (4) are closing the gap in a particularly rapid fashion, having already overtaken boys in terms of their literacy scores. The qualitative data at both baseline and midline suggested that the aggregate score-gap between boys and girls is in part a result of the fact that girls are less confident than boys in terms of participating in class and asking for help when they need it. It is thus important to consider that, as yet, there is no measurable intervention-effect on girls' life-skills. Life-skills is a major point where there is potential for gender transformative impact but where progress has not yet been made. At least in part, the problem may be that teachers continue to reinforce the gender gap by treating girls differently from boys in ways that may discourage girls in the classroom.

As at the baseline, boys and girls demonstrate the same fundamental skill gaps in terms of numeracy and literacy, which suggests that these gaps are probably a product of objective gaps in teachers' skills. In terms of attendance, there are also no significant differences between boys and girls.

Midline analysis reaffirms the fact that girls belonging to pastoralist households are among the most consistently marginalized. This finding is consistent across learning, transition, and attendance outcomes, suggesting that interventions specifically targeted at removing barriers for pastoralist households are needed.

At baseline, a number of project interventions were identified as having the potential to be gender transformative, including extensive engagement with the MoE, and the creation of a "broad social movement towards changes in gender norms and power relationships affecting girls (and also boys), using the VSLAs and literacy courses for mothers as platforms for dialogues on gender and girls' education." These interventions retain their potentially transformative nature at midline, but little measurable progress has been made along these lines thus far. As noted elsewhere, a simple reason for the lack of progress so far is the relatively short duration of implementation prior to conducting midline data collection.

Girls' and Boys' empowerment forums also have gender-transformative potential and GEFs have already shown signs of possibly contributing to improved learning and transition outcomes. Based on the project Theory of Change, the presumed mechanism for GEFs to affect learning and transition would be because GEFs focus on educating and re-educating girls and boys into more gender-equitable ways of thinking. This, in turn, should empower girls in ways that would be reflected in increased confidence and YLI scores, and then also be reflected in better participation and learning in the classroom, and which would also be reflected in more control over their life choices and the ability to stay in school and transition successfully. However, YLI scores have not increased measurably for girls in intervention schools or for girls who specifically participated in GEFs, which suggests that the correlation between GEFs and increased learning outcomes may be spurious or that the mechanism for improvement may not be the confidence-based mechanism hypothesised in the TOC. Unfortunately, the qualitative data does not address this particular question (of why YLI scores have not increased significantly for GEF participants), and future qualitative work would ideally do more to directly investigate this relationship.

In the qualitative data, there was some evidence that the project should continue to monitor and remain sensitive to the effects of project activities on boys. Girls were still commonly described as shy, but many respondents highlighted that girls are now outperforming boys in school, and in some communities,



respondents voiced concerns that girls' enrolment has surpassed boys enrolment. Additionally, boys in some cases appear to be less motivated than girls in school and are not treated with the same sensitivity that girls are – for example, teachers are willing to punish boys more harshly than girls. Misinterpretation of pro-girls' education messaging may be partially responsible for this lack of motivation observed among boys, as it seems teachers in some cases are motivating girls at the expense of boys – for example, highlighting how girls are needed to support their families because boys are less focused on their education.

## 8.2. Recommendations

### Monitoring, evaluation and learning of the project

- Skill gaps in learning have moved in a positive direction since the baseline, but advances in literacy skills have been very slow, there has been no detectable effect of the intervention on girls' learning, and teaching quality still ranks as one of the most important barriers at midline. Even by the most direct measures such as classroom observation, teaching quality has not improved measurably since the baseline, suggesting that efforts to improve teaching quality need to be redoubled in order for learning to improve. In addition, numeracy boost training is not predictive of improved numeracy outcomes among girls whose teachers received that training. All of these findings should be understood in light of the fact that project implementation has only occurred for a duration of four months, and thus there has been very little time for intervention effects to take hold in a way that would be measurable and detectable at the aggregate level. Thus, the findings related to learning suggest that teaching quality remains one of the most critical determinants of learning and should be monitored closely going forward. However, it should not be concluded that teaching quality is not responding to project interventions; rather, it is most likely the case that even if teachers have experienced immediate improvements as a result of training, students have not yet had time to fully benefit from improved teaching to a degree that would produce statistically significant improvements over baseline learning assessment scores. Having a more direct measure of teachers' subject-specific skills would allow for a more nuanced examination of this causal process. At the baseline it was suggested that a direct test of teachers' skills (through numeracy and literacy assessments) would be an ideal way of confirming if learners' skill gaps are a result of teachers' skill gaps, and this suggestion still applies.
- At midline, there is an apparent disconnect between interventions, YLI scores, and outcomes. As noted above, GEF participation is associated with higher learning and transition outcomes, but not with higher YLI scores. The result of apparently broken causal chain is a lack of clarity about the mechanism through which GEFs may be having an effect, as well as the more fundamental question of whether or not GEFs are indeed having their intended effect. Future qualitative data collection should focus on excavating the link between GEF programming, girls' confidence levels (including the specific measure of YLI scores), and girls behaviours such as classroom participation that might plausibly link interventions to outcomes via intermediate outcomes.
- In light of the qualitative and quantitative data suggesting that conflict had an important negative effect on learning and other key outcomes, a greater degree of micro-level analysis of the effects of conflict would be desirable. The current household survey does not contain any measures of whether or not a given household or community was affected by armed conflict, which makes it difficult to perform the desired analysis. This problem can be addressed through the future addition to the household survey of a question or questions asking the primary caregiver to report

whether or not their community was directly affected by armed conflict within the past year (as well as some measure of the intensity of the effect).

- Analysis of the teaching quality data revealed a greater need to differentiate between treatment of girls and treatment of boys. Qualitative data of questions relating to teaching quality revealed that girls are often not subject to as much corporal punishment as boys in the classroom. In addition, nearly half of girls in the midline, 44.8 percent, agreed with the statement that teachers treat girls different from boys, but it is not evident whether boys receive favourable treatment, whether girls receive favourable treatment, or whether the difference in treatment involves favourability of treatment at all. As such, we recommend that questions to girls about corporal punishment in the classroom are followed up with questions about whether boys, girls, or both were beat and questions about different treatment between boys and girls are followed up with questions about in what way the treatment was different.

### Project design and relevance

- Girls belonging to pastoralist households remain among the most marginal in the sampled population. The project has encouraged and facilitated CEC outreach efforts, but girls who belong to pastoralist households still have lower than average learning scores and higher than average absentee rates. Financial hardship is the primary mechanism through which drought affects the education of girls in pastoralist households. Notwithstanding the projects activities and CEC activism aimed at helping pastoralist families enrol their children and keep them in school, pastoralist families who have lost their herds (i.e. their main source of livelihood) because of drought may be unable to afford school fees or may lose a potential income-earner for the family by sending their children to school. Girls belonging to pastoralist households remain at high risk of missing days of school, and the economic distress in their households may reduce their study time and even their food security in ways that significantly affect their ability to learn in school. In particular, pastoralist girls who are re-enrolling in formal education will require additional tutoring or help with their studies in order to compensate for the fact that many have been out of school or have had lower-than-average attendance levels. Bursary packages may need to be adjusted to account for higher levels of economic hardship among pastoralist girls who are in school, but who face more barriers to studying and attendance than do their peers. Additionally, school feeding programs could also be explored as a way to reduce food insecurity among pastoralist children. There is evidence that CECs and parents are currently mobilizing to establish these programs, suggesting that there is a gap in this type of programming.
- Subgroup analysis of learning outcomes suggests that project efforts may have been successful in beginning to reduce the degree to which poverty is a barrier to learning. At the midline, girls belonging to households with proxies of economic distress still have lower than average learning outcomes, but not to a statistically significant degree. This is a quantitative improvement over the baseline findings, and it is also consistent with qualitative data suggesting that CEC engagement and other project efforts have helped to mitigate the influence of poverty on learning. However, our recommendation is that efforts to address poverty be redoubled, partly because they appear to have had a positive impact so far, and partly because there is still clearly room for progress as learning scores are still somewhat lower on average for poorer households, and poverty is a significant predictor of lower transition rates. Additionally, there should be an examination of project messaging and its interpretation within communities to ensure that educators and community members are not encouraging and motivating girls who are doing well in school at the exclusion of girls who are not doing well or of boys.

- Changes between baseline and midline coding of disability status (and the exclusion of OOS girls from the midline sample) have impeded longitudinal analysis of disability as a barrier. Nonetheless, a finding that remains consistent between baseline and midline, irrespective of the coding scheme used, is the fact that girls with reported vision impairments have significantly lower learning outcomes at baseline and midline. The consistency of this finding suggests that girls with vision disabilities are probably at a significant disadvantage in terms of learning vis-a-vis their peers, and thus the project should continue to target teacher training (and possibly other interventions as well) toward the accommodation of students with vision impairments. More broadly, there should be an increased focus on educating teachers and CEC members on the unique challenges faced by girls with disabilities. Training should enable teachers and CEC members to incorporate sensitivity to children with disabilities into their community outreach activities and encourage them to focus some fundraising efforts on meeting the needs of children with disabilities (e.g. raising funds for crutches, which would help a girl who has a physical disability reach school more easily). Teachers should also be sensitized on how to provide one-on-one support to children with disabilities in school, perhaps through the development, in coordination with parents, of specialized education plans for children with disabilities that addresses important topics for consideration, including: determining distance to school and what assistance is required, if any, to reach school; determining levels of access to treatment; determining barriers to attendance and how to best alleviate these barriers; ensuring the child does not fall behind in the event of absence; facilitating the child's learning in the classroom in light of their specific disability; monitoring inclusion and safety; providing one-on-one support after class, if needed; enacting a buddy system in class, if needed.
- There is also a very clear need for child protection policies that include plans for handling abuse directed toward girls, particularly girls with disabilities. Teachers should be equipped with skills and support that will allow them to provide students with a safe learning environment without feeling they will be targeted in retaliation. Additionally, CECs and school administration should be encouraged to develop formalized child protection policy plans that address the protocol for handling incidences of violence at school, including case referral, case escalation to actors outside the school administration, and establishment of clear child protection roles within the school administration. To supplement these efforts, CEC members and teachers should be supported to incorporate content into their outreach that addresses negative/harmful community attitudes toward girls, particularly girls with disabilities.
- Qualitative evidence suggests that some tensions exist between the head teachers and CEC members in some locations. Head teachers appear to complain about CEC members' lack of experience and exposure to training which may imply that the CECs may not have an adequate amount of skills in order to participate in school management effectively. Therefore, it may be necessary for the CARE International to conduct different/additional trainings in the areas of financial management, fundraising, awareness, and empowering girls in order to help CECs better understand their roles and responsibilities and engage more meaningfully in supporting girls' education and sustaining schools. From a sustainability perspective, it will also be crucial to ensure that the MoEs acquire the proper institutional capacity and skills to continue handling such trainings when the project ends.

### **Scalability and sustainability**

- The collection of attendance and enrolment records has improved dramatically from baseline to midline, however substantial gaps remain. Approximately 30 percent of cohort girls are in classrooms in which enrolment records for the school thus far were not kept. The effort to further

improve attendance can be supported by providing training and resources that will help teachers maintain attendance records. Furthermore, additional interventions should help teachers, principals, and other stakeholders use attendance record to identify and target interventions to at-risk girls.

# Annexes

## Annex 1: Midline Evaluation Submission Process

Please submit all Midline reports and accompanying annexes via Teamspace, an online file-sharing platform. Both the External Evaluator (EE) and Project should have access to their respective Teamspace folders, however please reach out to your EO if you do not.

Please note, Annexes can be uploaded to Teamspace for FM review separately and before the midline report analysis is completed. We advise Projects and EEs to follow the sequence outlined below to speed up the review process and avoid unnecessary back and forth. Where possible, we also advise that projects and EEs do not begin their ML report analysis until Annex 13 is signed off by the FM.

### **Annexes to submit for FM review any time before the ML report is completed:**

- Annex 2: Intervention roll-out dates.
- Annex 3: Evaluation approach and methodology.
- Annex 4: Characteristics and barriers.
- Annex 7: Project design and interventions.
- Annex 9: Beneficiaries tables.
- Annex 10: MEL Framework.
- Annex 11: External Evaluator's Inception Report (where applicable).
- Annex 12: Data collection tools used for midline.
- Annex 13: Datasets, codebooks and programs.
- Annex 14: Learning test pilot and calibration.
- Annex 15: Sampling Framework.
- Annex 16: External Evaluator declaration.
- Annex 17: Project Management Response (this can be revisited following feedback from the FM).

### **Annexes to finalise after Annex 11 "Datasets, codebooks and programs" is signed off by the FM:**

- Annex 5: Logframe.
- Annex 6: Outcomes Spreadsheet.
- Annex 8: Key findings on Output Indicators.

## Annex 2: Intervention roll-out dates

Please provide a timeline of roll-out of your interventions in the table below.

Activities	Start	End
<b>Output 1: Improving access to post-primary options</b>		
Meetings with MoEs, specialists and other stakeholders to develop ALP model	October, 2017	December, 2017
MoE subject specialist workshop to develop ALP modules	January, 2018	February, 2018
Validation, translation, production and distribution of ALP modules	July, 2018	September, 2018
Roll out of ALP classes (ALP implementation)	September, 2018	Ongoing - end date October, 2021
Develop girls' life skills in upper primary through ALP, including leadership skills, financial literacy and business selection and management of income generation activities; participation in Girls' Empowerment For a	July, 2018	Ongoing - end date October, 2021
Training of CECs to improve retention and transition ( 33 additional secondary schools)	February, 2018	July, 2018
CEC Coaching on improving retention and transition (199 schools)	March, 2018	Ongoing – end date September, 2020
Provide partial grants to girls from poor families	November, 2017	Ongoing – end date March, 2020
Equip and enrol girls into boarding schools	Not started	N/A
<b>Output 2: Supportive school practices and conditions for marginalised girls</b>		
Train teachers on improved delivery of literacy and English language, supported by digital content in all 148 primary and 55 secondary schools	February, 2019	Ongoing – end date May, 2019
Recruitment of consultant to develop manual and train teachers on improved delivery of numeracy	October, 2017	December, 2017
Refresher and advance numeracy TOT training	January, 2018	March, 2018
Train teachers on improved delivery of numeracy in all 148 primary and 55 secondary schools (cluster training)	July, 2018	September, 2018
Train teachers to provide structured remedial support to students at primary and secondary level	July 2018	Ongoing - end date October 2021
Train and coach teachers to deliver the ALP curriculum	July 2018	Ongoing - end date October 2021
Construct additional classrooms in remote primary schools; build water facilities in new secondary schools; and provide solar chargers for mobile devices/tablets and sanitary pads to schools	April, 2018	March, 2019

Activities	Start	End
Incorporate life skills and financial literacy training into GEFs and BEFs	April 2019	Ongoing – end date October 2021
Provide career guidance in schools	November 2018	Ongoing - end date October 2021
<b>Output 3: Positive shifts on gender and social norms at community and individual girl level</b>		
Engage community-level stakeholders including religious leaders, women’s groups, men and boys	February, 2017	September, 2018
Expand and strengthen GEFs and create BEFs to develop leadership and mentorship skills	September, 2018	Ongoing – end date: March, 2019.
Provide adult literacy and financial literacy classes for mothers	May, 2018	March, 2019
Support the financial empowerment of mothers through savings groups (VSLA), business selection, and business coaching and mentoring	February, 2018	Ongoing – end June, 2019
<b>Output 4: Enhanced MoEs’ capacity to deliver quality and relevant formal and informal education</b>		
Strengthen Gender Departments’ capacity to improve girls’ education outcomes through trainings, development of action planning and provision of incentives to retain the gender focal points especially in rural areas	December, 2017	Ongoing - end date Dec. 2019
Support quality assurance and standards (QAS) functions at all MoE levels	September, 2018	Ongoing – end date June 2020.
Provide support to Regional Education Officers (REOs) and District Education Officers (DEOs) to mainstream improved teaching practices and address retention/ transition	January, 2018	Ongoing – end date June, 2019
Work closely with MoE on NFE for mothers and entrepreneurship skills for girls	April, 2018	December, 2018
Development of project IEC materials in conjunction with MoE for use at stakeholder advocacy and promotion events	July, 2018	Ongoing – end date March, 2021

## Annex 3: Midline evaluation approach and methodology

### Evaluation Design

The midline evaluation of SOMGEP-T builds on the baseline, whose data collection was conducted from late October to early December, 2017. The project’s MEL Framework specified the use of a quasi-experimental research design, which the midline evaluation uses. The quasi-experimental design uses an explicit comparison group for analysis in a difference-in-differences framework.



The quasi-experimental design was selected because it best balances the needs of rigor in the evaluation and the logistical, ethical, and practical difficulties of actually conducting an evaluation. In other words, the quasi-experimental design is a rigorous approach – under a set of well-known and frequently plausible assumptions, valid causal inferences can be drawn about a project's impact. The design relies on two assumptions: the first is often called the “parallel trends assumption,” which states that outcomes in the intervention and comparison groups, in the absence of the intervention, would have evolved similarly over time (i.e. their trends would have been parallel from baseline to endline if no intervention had occurred). This is often referred to as considering the counterfactual – what would have occurred in the intervention group in the absence of the intervention, and whether it would have mirrored the actual evolution of the comparison group. The second assumption is that intervention effects do not spill over from intervention to comparison units. To the extent that the two groups are cross-contaminated, often, but not exclusively, through migration, it invalidates the inferences drawn about project impact.

Critically, if these two assumptions are met, the quasi-experimental design allows valid *causal* inferences about project impact. While these assumptions are more stringent and harder to satisfy than those required in a randomized controlled trial (RCT), the quasi-experimental has advantages over the RCT in some cases, especially those where it would be difficult to convince control communities to participate in an RCT. Many observers have expressed ethical concerns about the use of RCTs, because individuals are asked to participate in research without the opportunity to benefit from the project being studied. More importantly, research subjects occasionally object along similar lines, which can make it challenging for organizations to conduct an RCT and simultaneously maintain positive relations with communities where they may wish to work in the future.

The baseline evaluation utilized a joint sampling approach, in which one cohort of girls was selected to serve as both learning and transition cohorts. Baseline sampling occurred in two stages: first, schools were selected after matching intervention and comparison schools using coarsened exact matching. Selection with matching ensured that the intervention and comparison schools were balanced on a small set of characteristics, such as support from other NGOs, that would be expected to influence the targeted project outcomes. Once a balanced sample of intervention and comparison schools was chosen, the selection of the cohort girls consisted of a fully random household survey in the selected communities. Households were selected using varied starting points and a random walk methodology, and households were screened for eligibility, depending on whether their household included a girl in the target age range. The resulting sample has the advantage of being representative of the communities in which the evaluation is being implemented (as opposed to sampling from among girls at school, which systematically excludes out-of-school girls). It

Section 2.2 details changes made to the sampling design at the midline and we quote extensively from it here to make clear the decisions that went into the sample design in this round:

*Two important changes to the sample design were made at the midline. The first concerns the sample of schools which will be visited; the second concerns the set of girls in the baseline cohort who will be re-contacted.*

*At the midline, the evaluation team visited a truncated set of the same intervention and comparison schools from the baseline. At the baseline, 76 schools were visited in total, with the sample evenly split between intervention and comparison schools, which were matched using Coarsened Exact Matching (CEM) to create a sample balanced on pre-intervention observable characteristics, to the extent possible.*

*At the midline, the evaluation team will re-visit 63 of the same schools. Schools were removed from this round of data collection for both security and logistical reasons.*

*Notably, five schools were excluded from the midline sample because they were outliers – in terms of learning outcomes, especially – at the baseline. The five schools had particularly skilled teachers with regard to English-language instruction, and were typically removed from consideration during the analysis of the baseline data (or different sets of results – some including them and some excluding them – were reported). Note that the excluded schools can be reincorporated into the sample in future evaluation waves.*

*The second sampling adjustment concerns the set of cohort girls who were re-contacted at the midline. At included schools, the evaluation team re-contacted a subset of the girls learning and transition cohort (i.e. the baseline sample of girls). Specifically, the midline sought to re-contact only girls who were enrolled in school at the time of the baseline. Girls who were out-of-school at the baseline – whether enrolled in alternative education or employed or not otherwise engaged – were not re-contacted in this round of data collection. Importantly, all members of the learning and transition cohort will be re-contacted at the third evaluation round and endline for the purposes of assessing aggregate transition and learning outcomes over the life of the project.*

*Naturally, these sampling decisions have important consequences for the types of analysis that can be conducted in this round. First, all analysis that makes comparisons between baseline and endline – including comparisons of learning and intermediate outcomes – uses the subset of respondents or observations in the baseline that are comparable to the midline. This approach is simplest in the case of school- or classroom-level analysis, such as the analysis of the head teacher survey, headcounts, or classroom observations. In these cases, we subset the baseline sample to include only those schools that also appear in the midline sample. The population being studied has, therefore, changed: when we assess a change in attendance rates via classroom headcounts, we can only say that the change in attendance rates from baseline to this second evaluation round applies to the subsample of schools that was included in both waves.*

Given the complexities introduced by the sampling decisions – and the fact that these complexities are often context- or outcome-specific, we review sampling considerations at length throughout the report. For instance, in our analysis of transition outcomes, we discuss how comparable baseline and midline transition subsamples were constructed from the available data, and the manner in which our analysis is influenced by the nature of those samples.

The evaluation employed a mixed-method approach, collecting both quantitative and qualitative data from a wide range of beneficiary and stakeholder populations. Quantitative data at the midline was collected, separately, from school-age girls, their caregivers, their head of household, head teachers or principals, and teachers via direct classroom observation. Qualitative data was collected from girls, mothers, teachers, CEC members, and others.

Qualitative data was used to contextualize key quantitative findings and to explore the context of those findings. However, it was also treated separately from the quantitative findings, in an effort to “let the data speak for itself.” A dedicated qualitative analyst read through the entirety of the qualitative interviews – without having previously conducted analysis of the quantitative data – allowing themes to emerge from the interviews naturally. Only after reviewing the interview transcripts and writing up initial findings did the qualitative analyst read findings generated by the quantitative analysts, to ensure that qualitative findings were not affected by the results of the quantitative analysis. The findings were then incorporated into the

report wherever appropriate – as stand-alone findings on indicators with limited quantitative data, as stand-alone findings on indicators where the quantitative data did not address all aspects of the indicator, as a source for triangulation of other findings, and as contextualization of the quantitative results.

### Prior to Data Collection

Relatively few changes were made to the evaluation design prior to the start of data collection, and data collection generally followed the original evaluation plan developed in CARE's MEL Framework. As noted above, the evaluation team sought to re-contact at midline those cohort girls who were enrolled in school at the baseline. This criterion, and the selection of schools to be visited at the midline, yielded a set of 828 girls to be targeted for re-contact. The final achieved sample was 807 girls. The achieved sample fell below the target sample for two reasons: either they were contacted – and even interviewed – but were determined to have been ineligible for inclusion in the target sample from the start (typically because they fell outside the target age range or were not truly enrolled at the baseline, due to errors or deliberate misrepresentation at the baseline); or girls could not be located and, when a replacement girl was sought, there were insufficient girls in the school to allow for replacement.

Of course, a number of girls were unsuccessfully re-contacted at the midline, and were replaced with girls at the same schools. The re-contact and replacement process was spelled out explicitly for team leaders and enumerators, to avoid unnecessary panel attrition. Before replacing a girl, the following actions were taken to re-contact her or her family:

- Check at the school for the girl, including with the head teacher and teachers
- Ask girls in her same age cohort whether they know her and where she can be located
- Call the household's primary and secondary phone number – given at the baseline – to locate her
- At least two contact attempts for each available phone number were made, spread at least 6 hours apart
- Visit the household to locate the girl
- At least two visits were made to the household, at least six hours apart
- Contact the household's neighbours
- Ask about the girl and her family at local meeting points, such as the community's mosque

If a girl's household was reached and they confirmed that the girl was unavailable, or the field team completed all of these steps and could not locate the girl, she was replaced at the same school. Replacement girls were selected from the set of enrolled girls, attempting to match the age and grade level of the replaced girl as closely as possible.

Drawing from Section 2.4 of the report:

*Overall, the re-contact rate in the sample was 80.3 percent, with 159 cohort girls of 807 replaced at midline. In intervention areas, 79.6 percent of girls were re-contacted successfully, compared to 81.1 percent of girls in comparison areas ( $p = 0.59$ ).*

*The SOMGEP-T sampling strategy was developed with a high level of panel attrition in mind. In the project's MEL Framework, the sampling calculations included a buffer for up to 40 percent attrition. At this stage, the evaluation is just within the expected range of attrition, at 19.7 percent from baseline to midline. An important open question concerns how attrition rates will differ from the 2018 midline to the 2019 second midline and from the 2019 midline to endline. Because the midline only sought to re-contact girls who were in school at the baseline, attrition rates may be higher among the full sample of cohort girls and*

*go beyond the assumption of 40 percent.<sup>328</sup> On the other hand, to the extent that the project itself improves transition rates, panel attrition may decline over time, at least in intervention areas. Of course, a number of unknown factors will influence attrition rates over the next two years, including rainfall patterns, migration, and conflict dynamics. At the midline, panel attrition was high, but manageable; additional efforts will need to be made at the endline to increase successful re-contacts as much as possible.*

Data collection tools were not changed dramatically from the baseline to midline. The only changes to the quantitative tools were:

- Minor adjustments to the learning assessments, while maintaining their approximate difficulty (see Annex 14 for discussion and analysis)
- Removal of the self-administered teacher survey
- Minor expansions of the classroom observation and head teacher survey tools to address specific areas of interest

Qualitative tools underwent more significant revision at the midline, including the introduction of two entirely new interview types. FGDs with boys and girls, separately, were employed at the baseline; in lieu of these interviews, the midline included two types of participatory group discussions with girls: a risk mapping exercise, and a story-telling or vignette exercise. In the latter type of interview, girls were introductory vignettes describing a girl “living in a community like this one.” The girls were asked to discuss in a group how the story would turn out, reflecting on the barriers that “girls like them” face to completing their education, the types of support they would expect to receive, and so forth. The girls were asked to complete the story in the way that they thought was most likely; the qualitative researcher probed for additional detail, asked them to consider alternatives, asked them why an alternative might not happen or rule out alternative endings to the story. Four risk mapping and four vignette exercises were completed, in total, with different groups of girls.

The data collection teams consisted of a mix of Forcier’s core national staff members and experienced enumerators from a roster of researchers that Forcier has employed on previous projects. Eight data collection teams were utilized, each headed by a team leader. Four of the eight team leaders were members of Forcier’s core staff, i.e. permanent, salaried researchers in Forcier’s Hargeisa office. The remaining four team leaders were experienced researchers that Forcier has employed extensively in the past. Three of the team leaders were also team leaders during the baseline evaluation; the remaining five team leaders had served as team leaders on previous GEC evaluations in Somalia (i.e. evaluations of Relief International’s EGEP and EGEP-T projects). Nearly every enumerator employed on the evaluation had participated in fieldwork on at least one prior GEC or GEC-T project in Somalia, with some working on their fourth or fifth such engagement.

Training took place over the course of seven days and was led jointly by a team of technical staff from Forcier, joined by CARE’s monitoring and evaluation team.<sup>329</sup> The first day introduced the project, provided an overview of the tools and fieldwork process, and covered child protection policies. Days 2 and 3 were spent reviewing the tools in detail as a group, verifying the quality of translations, and ensuring that all enumerators understood the content of the questions and the filter logic of the script.

---

<sup>328</sup> This would occur if OOS girls were less likely to be successfully re-contacted later – if OOS girls are more likely to migrate, get married, or have less stable households, this may be the case.

<sup>329</sup> Forcier’s training team consisted of their Director of Research for the Horn of Africa, their Director of Training, and the Deputy Country Director for Somalia and Somaliland. Modules of the training were led directly by the most experienced team leaders, so that it could be conducted comfortably in Somali.

Days 4 and 5 were primarily spent practicing the household survey and learning assessments in group and one-on-one settings. Day 6 consisted of a pilot of the household survey and learning assessment. Day 7 was dedicated to debriefing and a review of lessons learnt from the pilot, final guidance for fieldwork, and distribution of fieldwork materials to the teams and team leaders.

Qualitative research was typically conducted by the team leaders, who all have extensive experience conducting FGDs and KIIs on education-related projects, including GEC evaluations. The exception consists of interviews with girls (girls with disabilities, risk mapping, and vignette exercises), when the team leader in question was male. In these cases, the most experienced female team member was selected to conduct the qualitative interview. In general, the evaluation team sought to focus the two participatory exercises – which were new to the team leaders and considerably more difficult than typical FGDs – in areas to be visited by female team leaders and by the most experienced team leaders. As noted above, most of the team leaders conducting qualitative interviews had previously served as team leaders – and therefore conducted similar qualitative interviews – for previous GEC evaluations.

Qualitative researchers underwent separate training while enumerators were practicing the household survey and learning assessments. Specifically, the qualitative tools were reviewed as a group, with an open discussion regarding the meaning and purpose of each question. This process also allowed the group to check the quality of the translations, though this was not the primary focus of the session; rather, gaining an understanding of the purpose of the research, to facilitate more active and effective probing, was the goal of the training sessions. Particular time and energy was spent introducing the participatory exercises, because they were unfamiliar to the qualitative researchers. In total, training of the qualitative tools took the bulk of two days of training.

### **During Data Collection**

Data collection took place from November 6 to December 10, 2018. Fieldwork in all regions began simultaneously, immediately following training, though the ending dates of fieldwork varied from region to region, depending on the sample size of schools in a given area. Because enumerators and team leaders must be drawn – for security and other reasons – from the region in which they work, it is not typically possible to re-allocate teams to other areas when they finish fieldwork. Therefore, regions with smaller sample sizes in terms of either schools or girls to be contacted, finished fieldwork earlier. Qualitative data collection was not staggered; visits were structured such that qualitative and quantitative data collection took place at the same time, during visits to a community ranging from 1 to 3 days in length.

The field teams were trained on and followed child protection standards set out by Forcier as part of its standard employment contract and by CARE International as a separate policy they require of their evaluators. Security concerns were handled on a case-by-case basis, with Forcier’s Somalia-based operations staff making determinations regarding the safety and accessibility of particular areas where necessary. None of the 63 sampled schools were excluded or substituted for security reasons, though the construction of the midline sample itself did exclude two schools from the baseline sample due to security concerns (as noted in the sampling discussion presented in the main report).

The sampling methodology is described in detail in Section 2.2 of the report, and the re-contact and replacement protocols for cohort girls are described in the section above, which provides the full procedure field teams employed for contacting a girl and subsequently replacing her. The precise sampling procedure did vary slightly from tool to tool, given that different populations were targeted. A census was conducted among head teachers, who completed the school survey, as the head teacher at every sampled school was interviewed. A virtual census was also completed for the headcount tool: every

eligible classroom from grades 1 through 8 was surveyed for the headcount tool. Classroom observations, in contrast, were conducted in just 1 classroom per school; the classroom was randomly selected from among those where the students were learning Somali, English, or mathematics. The sampling protocol for girls and, by extension, their households, is described in Section 2.2.

During fieldwork, the data underwent significant quality assurance and quality control. Prior to the start of fieldwork, the evaluation team developed a set of quality checks, which supplemented the team's standard quality control script and team leaders were trained to conduct in-person quality assurance.

In the field, team leaders accompanied enumerators on a set number of interviews, to verify the quality of their work. Team leaders also conducted random callbacks for the same purpose.

On a daily basis, Forcier's Research Officer downloaded the data from Ona's server and ran the quality control script; a representative but not exhaustive list of the daily checks includes:

- Long and short duration surveys
- Assessment of the composition of key demographic variables, to ensure they are in line with expectations (e.g., grade and age distribution)
- Surveys with unlikely GPS locations or GPS locations that change during the interview
- Logical consistency checks among related questions (e.g., word-per-minute scores among children who could not identify letters, etc.)
- Checking response or re-contact rates
- Identifying surveys with non-response or large numbers of "don't know" responses
- Checking for logical consistency between midline and baseline responses (logical grade progression; stability in demographic variables, etc.)

In each case, the Research Officer identified observations that the script flagged, checked them manually, contacted the field teams to verify the information, and made corrections to the data in a canonical and replicable cleaning script (.do file). On a less frequent basis, Forcier's Global Technical Team undertook additional, more extensive quality control checks, such as checking for "enumerator effects" in learning scores and other key outcomes. Both sets of checks typically required a number of callbacks or follow-up visits to households to verify or correct the data collected.

The final sample sizes achieved for each data collection tool are provided in the table below.



Table 89: Tool Details

Tool (used for which outcome and IO indicator)	Beneficiary group	Actual sample size intervention and (control group) - if appropriate	Remarks: 1) Attrition rate from baseline to midline 2) Re-contacted sample vs replaced sample 3) Major changes to tools or differences between anticipated and actual sample sizes
SeGMA and SeGRA (for learning outcome)	In-school and out-of-school cohort girls (aged 10-19 years at baseline)	807 (431 intervention; 376 comparison)	Targeted only girls who were in-school at the baseline. Identified 828 girls who fit those criteria, according to data from the baseline. Achieved sample of 807 girls, excluding girls who either did not actually meet the criteria based on new information collected at midline, or who needed to be replaced but could not be due to a small number of girls in their respective schools.  648 of 807 (80.3%) girls successfully re-contacted and interviewed. 159 (19.7%) girls replaced with similar girls, drawn from the same schools (all, by definition, enrolled in school at midline)
Household survey	In-school and out-of-school cohort girls and their households	807 (431 intervention; 376 comparison)	Same sampling notes as SeGMA and SeGRA, above (same sample composition and replacement rates)
Head Teacher (School) Survey (for school management and other outcomes)	All head teachers at sampled schools	63 (32 intervention, 31 comparison)	
Headcount Tool (for attendance outcome)	All classrooms at sampled schools	455 (232 intervention, 223 comparison)	Census of classrooms in grade 1-8 at sampled schools. Note that combined classrooms – i.e. those with grade 7 and 8 students – were recorded separately as two observations with distinct headcounts and attendance rates.
Classroom observation (for teaching quality outcome)	1 classroom per school	62 (32 intervention, 30 comparison)	<ul style="list-style-type: none"> <li>• One school excluded from sample</li> <li>• Targeted classes were teaching Somali, English or mathematics</li> </ul>



Following the conclusion of data collection, the data were cleaned and checked for consistency, including a more intensive set of quality checks. This process identified a small number of issues that were not already identified during quality control checks employed during the fieldwork period. In these cases, team leaders performed callbacks to verify information from households and head teachers, as needed.

Throughout data collection, there was constant communication between team leaders and enumerators in the field, and the Research Officer in the national office. This communication allowed the Research Officer to keep track of issues that arose during fieldwork, and record the researchers' insights and experiences. At the conclusion of fieldwork, team leaders were debriefed and their observations recorded.

During fieldwork, the quantitative data was uploaded daily – or else whenever network service would allow – to Ona's server. The data was downloaded, checked, and basic cleaning tasks were performed daily. After fieldwork was over, the evaluation team conducted a more thorough cleaning, including merging the midline data with the baseline data for consistency checks and to prepare for analysis. All cleaning and quality control actions were performed in Stata, using scripts (.do files) to facilitate replication.

All qualitative interviews were audio recorded. Qualitative researchers took notes immediately after each interview, to ensure that their recollections were fresh. Audio files were returned to the national office, where experienced staff transcribed the audio recordings and, separately, translated, verbatim, the transcriptions into English. Each transcription and translation was checked by a second Somali member of Forcier's core national staff, to verify its quality. A final check was performed by one of Forcier's Research Officers or Research Managers, during which they reviewed the English translations for coherence; translations with probable mistakes were returned to the translator and staff member who originally performed its quality check. The English translations were provided to the evaluation team's dedicated qualitative analyst. As discussed in our description of the evaluation design, the qualitative and quantitative analysis were kept strictly separate at the start, to allow themes to emerge naturally from the qualitative data. Only after thoroughly reviewing the interview transcripts and writing a first set of results on that basis did the qualitative analyst begin to review the quantitative findings.

Beneficiary tracking in the next evaluation wave will proceed exactly as it did from the baseline to midline. Each sampled girl and their household provided contact information in the form of one or two phone numbers; the locations of their household was recorded in verbal (i.e. written directions) form and a GPS location was recorded as well. In the next evaluation wave, teams will attempt re-contact using the same procedures outlined and followed at the midline.

### **Limitations and Challenges**

A list of limitations to the evaluation design – both in principle and in practice – is provided in Section 2.3 of the report.

## Annex 4: Characteristics and Barriers

The tables below describe the characteristics of the baseline and midline samples of cohort girls, respectively. Note that, to facilitate comparisons between the baseline and midline sample, we have limited the baseline sample to the portion that is comparable to the midline. As discussed in Section 2.2 of the report, midline data collection occurred in 63 of the 76 schools visited during the baseline, re-contacting only those girls who were enrolled in school at the baseline. The baseline sample analysed in this section are *all cohort girls* sampled at baseline from the 63 schools that were visited in both rounds. The midline results are drawn from the unabridged midline sample.

Table 90: Girls' characteristics

	Intervention (midline)	Control (midline)	Source (Household and Girls School survey)
<b>Sample breakdown (Girls)</b>			
<b>Orphans (%)</b>	Midline (baseline)	Midline (baseline)	
- Single orphans	13.7% (11.1%)	13.1% (9.8%)	singleorphan
- Double orphans	0% (0.3%)	0% (0%)	doubleorphan
<b>Living without both parents (%)</b>	9.3% (6.1%)	8.2% (5.9%)	wobothparents
<b>Living in female headed household (%)</b>	45.5% (42.6%)	44.9% (44.6%)	femalehh
<b>Married (%)</b>	1.2% (0.9%)	2% (1%)	married_s
<b>Mothers (%)</b>			
- Under 18	0.3% (0.3%)	1.3% (0%)	under18mother
- Under 16	0% (0.3%)	1% (0%)	under16mother
<b>Poor households (%)</b>			
- Difficult to afford for girl to go to school	12.9% (21.1%)	7.6% (18.1%)	notaffordschool
- Material of the roof (material to be defined by evaluator)			
- Gone to sleep hungry for many days in past year	36.7% (32.4%)	40% (37.7%)	poorroof
	9.1% (7.3%)	12.9% (8.6%)	hungrymanydays
<b>Parental education</b>			
- HoH has no education (%)	49.1% (67%)	51.2% (64%)	hohnoedu
- Primary caregiver has no education (%)	51.3% (67.6%)	59.7% (75.1%)	pcgnoedu

### Barriers

The table below reports the prevalence rate of various potential barriers to learning and transition, disaggregated by intervention status and round. As noted above, the baseline sample is limited to the same set of 63 schools visited at the midline.

Table 91: Potential barriers to learning and transition

	Intervention (Midline)	Control (Midline)	Source
<b>Sample breakdown (Girls)</b>			
<b>Home – community</b>			
<i>Safety:</i>			
Doesn't feel safe travelling to/from school (%)	Midline (baseline) 3.7% (3.2%)	Midline (baseline) 2.5% (9%)	unsafetravelgirl
Girls Travels more than 30 minutes to school (%)	5.2% (7.7%)	6.5% (4.1%)	long_trip
<i>Parental/caregiver support:</i>			
Doesn't get support to stay in school and do well (%)	1.9% (6.1%)	2.9% (1.1%)	nosupport
Girl has no choice in whether to attend school (%)	86.1% (86.3%)	93.5% (85.6%)	no_choice
<b>School level</b>			
<i>Attendance:</i>			
Attends school half the time (%)	1.8% (1.2%)	1.4% (2%)	halftime
Attends school less than half time (%)	0.3% (0.9%)	2.9% (0.7%)	lhalftime
Doesn't feel safe at school (%)	2.2% (4.9%)	1.4% (6.4%)	unsafeschool
<i>School facilities:</i>			
No seats for all students (%)	13.3% (20.9%)	15.8% (30.5%)	noseats
Doesn't use drinking water facilities (%)	10.5% (19.7%)	23.7% (32%)	nodrinkingwater
Doesn't use toilet at school (%)	17.9% (22.6%)	28.8% (30.1%)	notoilet
No computers in class (%)	95.7% (89.6%)	93.2% (95.1%)	no_computers
Cannot use books or other learning materials at school (%)	20.1% (19.9%)	28.3% (25.1%)	no_materials
<i>Teachers:</i>			
Disagrees teachers make them feel welcome (%)	4% (5.1%)	5.8% (7.5%)	nowelcome
Agrees teachers treat boys and girls differently in the classroom (%)	43.1% (41.12%)	46.5% (41.4%)	treatdiff
Agrees teachers often absent from class (%)	17.2% (35%)	29.2% (40.1%)	teacher_absent
Afraid of teacher (%)	75.9% (57.7%)	82.7% (58.3%)	teacher_afraid
Uncomfortable asking teachers question (%)	1.5% (5.5%)	2.5% (4.9%)	teacher_noquestion
Teacher punishes/disciplines when students get lesson wrong (%)	80.2% (76.8%)	84.9% (79%)	teachers_punish
Physical punishment witnessed last week (%)	54.5% (63%)	51.2% (49.6%)	corp_punish
Caregiver rates quality of teaching as poor (%)	1.6% (3.3%)	2.9% (4.8%)	teaching_poorquality

## Annex 5: Logframe

The project's logical framework is attached as a separate annex.

# Annex 6: Outcomes Spreadsheet

The Outcomes Spreadsheet is attached as a separate annex.

## Annex 7: Project design and intervention

### Project to complete

Complete the following table.

Intervention types	What is the intervention?	What output will the intervention contribute to?	What Intermediate Outcome will the intervention contribute to and how?	How will the intervention contribute to achieving the learning, transition and sustainability outcomes?
Access	Developing and implementing Alternative Learning Program for Out of School Girls	Output 1	IO- 1	By offering an alternative pathway for girls who may have otherwise dropped out, transition rates will improve. Girls will have increased exposure to higher learning, which will boost learning outcomes. ALP's particular focus on developing life skills will ensure this intervention produces sustainable outcomes, or outcomes that are relevant to the individual and community.
	Provision of partial grants to girls from poor families	Output 1	IO- 1	Increased attendance and retention is expected to improve transition rates and learning outcomes, as girls who are in school and are properly equipped are more likely to succeed. Girls from poor families who may not have otherwise had access to education will be better equipped to participate in decision-making and economic activities.
	Equip and enrol girls in 2 boarding schools	Output 1	IO- 1	
School Governance/management	Capacity building of CEC's to improve retention and transition	Output 1	IO- 2	A focus on retention and transition is expected to have a direct impact on transition rates and learning outcomes, as girls will have better access to higher education levels. The focus on the community level will ensure buy-in and contribute to the project's sustainability at the community level.
Teachers capacity building	Train teachers on improved delivery of literacy and English language supported by digital content	Output 2	IO- 3	Improved teaching quality contributes to enhance learning and transition outcomes, as children are equipped with the literacy skills in Somali, English, numeracy and life skills necessary to progress to higher levels of education. Interventions focused on improving teaching quality are expected to boost transition rates and learning

	Train teachers on improved delivery of numeracy	Output 2	IO- 3	outcomes in a sustainable way, by equipping children with the skills they need to succeed not only in school, but outside school as well
	Train teachers to provide structured remedial support to students at primary and secondary levels	Output 2	IO- 3	
	Train and coach teachers to deliver ALP curriculum	Output 2	IO- 3	
	Train and coach teachers on career guidance	Output 2	IO- 3 & IO- 4	Encouraging girls to think about their futures and how to achieve their aspirations will impress on them the importance of knowledge and education. It will also give them a clear pathway to achieving their goals.
Community-based attitudes and behaviour change	Engage community – level stakeholders including religious leaders, women’s groups, men and boys	Output 3	IO- 1	Boosts to attendance and retention are expected to contribute to improvements in transition and learning outcomes. Shifts in gender and social norms are expected to have a long-term, sustainable impact on the communities in which SOMGEP-T will operate.
	Provide adult literacy and financial classes for mothers	Output 3	IO- 1	
Economic empowerment	Support the financial empowerment of mothers through savings groups(VSLA) , business selection and business	Output 3	IO- 1	

	coaching and mentoring			
	Work closely with MoE on NFE for mothers and entrepreneurship skills for girls	Output 4	IO- 1	Enhancing the capacity of MoEs to take action on girls' education will have long-term effects on the communities in which SOMGEP-T operates. It will encourage positive shifts in gender and social norms, and will give MoEs actionable ways to contribute to improving learning and transition outcomes.
Life Skills	Develop girls life skills in upper primary through ALP	Output 1	IO- 4	The project's learning outcomes are focused on literacy, numeracy, and financial literacy. This intervention is designed to boost these specific learning outcomes, as well as increasing the likelihood of transition into ALP or secondary education. Additionally, the focus on leadership skills and other skills relevant to the job market contributes to the sustainability of SOMGEP-T.
	Incorporate life skills and financial literacy training into GEFs and BEFs	Output 1	IO- 4	Financial literacy training is one of the specific learning outcomes SOMGEP-T is expecting to influence. Financial literacy and life skills training will increase the likelihood of girls succeeding in higher levels of education, and will also equip them to contribute to the local economy through income-generating activities. These skills are expected to increase the relevance of education for students and families. Life skills – specifically leadership skills – are expected to boost students' voice and self-confidence, enhancing classroom participation among girls.
Girls self Esteem	Expand and strengthen GEF's and create BEFs to develop leadership and mentorship skills	Output 3	IO- 4	Girls who receive leadership and mentorship skills through life skills development will be better equipped to participate in class, breaking traditional norms that restrict girls' voice; to engage in the local economy; and to contribute to their communities in the future. Additionally, the capacity of GEFs and BEFs to track attendance and retention rates will contribute to improvements in learning and transition outcomes, and will encourage community-based organizations to think about how their actions have a direct effect on important student outcomes.



MoE Capacity building	Strengthen Gender Departments' capacity to improve girls' education outcomes through trainings, development of action planning and provision of incentives to retain the gender focal points especially in rural areas	Output 4	All IO	Enhancing the capacity of MoEs to take action on girls' education will have long-term effects on the communities in which SOMGEP-T operates. It will encourage positive shifts in gender and social norms, and will give MoEs actionable ways to contribute to improving learning and transition outcomes.
	Provide support to Regional Education Officers (REOs) and District Education Officers (DEOs) to mainstream improved teaching practices and address retention/transition	Output 4	All IO	Enhancing the capacity of MoEs to take action on girls' education will have long-term effects on the communities in which SOMGEP-T operates. It will encourage positive shifts in gender and social norms, and will give MoEs actionable ways to contribute to improving learning and transition outcomes.
Construction	Construct additional classrooms in remote primary schools; building water facilities in new secondary schools and [provide solar chargers for	Output 2	IO- 2	Boosts to attendance and retention are expected to contribute to improvements in transition and learning outcomes. Infrastructure development will benefit not just the current cohort of students with which SOMGEP-T is engaged, but will also benefit future students

	mobile devices devices/tablets and sanitary pads to schools			
--	---	--	--	--

## Annex 8: Key findings on Output Indicators

**This annex should be completed by the project.**

The Evaluator should hand over any output-related data to the project to enable the project to populate the following tables.

Fill in the table below with every Output Indicator, means of verification/sources, and the frequency of data collection. Please include output indicators for which data collection has not yet taken place and state when data collection for these will take place.

*Table 92: Output indicators*

Logframe Output Indicator	Means of verification/sources	Collection frequency
<b>Number and Indicator wording</b>	List all sources used.	E.g. monthly, quarterly, annually. NB: For indicators without data collection to date, please indicate when data collection will take place.
<b>Output 1: Improved access to post-primary options</b>		
<b>Output 1.1: Percentage of project locations with an alternative learning program for upper primary/ secondary</b>	ALP monitoring tool	Monthly
<b>Output 1.2: Percentage of ALP groups providing life skills training to marginalised girls</b>	ALP monitoring tool	Monthly
<b>Output 1.3: Percentage of girls receiving partial grants who remain in school</b>	Partial Grants Fidelity of Implementation	Termly
<b>Output 2: Supportive school practices and conditions for marginalised girls</b>		
<b>Output 2.1: Percentage of teachers not using corporal punishment</b>	Midline Evaluation, GEF/BEF Fidelity of Implementation	Yearly
<b>Output 2.2: Percentage of teachers using the digital learning platform</b>	Classroom Observations	Monthly
<b>Output 2.3 Percentage of Girls' Empowerment Forums providing life skills sessions according to the guidance</b>	GEF Fidelity of Implementation	Bi annual

<b>Output 3: Positive shifts on gender and social norms at community and individual girl level</b>		
<b>Output 3:1 Number of women mentors providing support to marginalised girls</b>	GEF Fidelity of Implementation , Monitoring Visits	Monthly
<b>Output 3:2 Number and percentage of mothers completing literacy courses</b>	NFE Completion records	Annual
<b>Output 3:3 Percentage of active village savings groups in project areas</b>	VSLA Fidelity of Implementation [ FOI]	Monthly
<b>Output 4: Enhanced MOEs' capacity to deliver quality and relevant formal and informal education</b>		
<b>Output 4:1 Number of Gender Units conducting activities to promote girls' transition and learning</b>	Gender Units Reports	Monthly
<b>Output 4:2 Percentage of REOs/ DEOs engaged in joint monitoring visits to formal schools/ ALP classes to support teachers</b>	Joint Monitoring Reports	Quarterly

Report on the midline values/midline status of each Output Indicator in the table below. Reflect on the relevancy of the Output Indicator for your Intermediate Outcomes and Outcomes and the wider Theory of Change based on the data collected so far. Are the indicators measuring the right things? What do the midline values/midline status mean for the implementation of your activities?

*Table 93: Midline status of output indicators*

Logframe Output Indicator	Midline status/midline values Relevance of the indicator for the project ToC	Midline status/midline values
<b>Number and Indicator wording</b>	What is the contribution of this indicator for the project ToC, IOs, and Outcomes? What does the midline value/status mean for your activities? Is the indicator measuring the right things? Should a revision be considered? Provide short narrative.	What is the midline value/status of this indicator? Provide short narrative.
<b>Output 1: Improved access to post-primary options</b>		
<b>Output 1.1: Percentage of project locations with an alternative learning program for upper primary/ secondary</b>	<p>The ALP sites established by the project offer out of school girls post primary opportunity. The 1332 girls enrolled in ALP are expected to have increased learning outcomes as well as developing essential life skills which will enable them to be productive members of the society.</p> <p>Constant measurement of the coverage of ALP is vital. The indicator is still relevant; no modifications are required.</p>	<p><b>Midline Wave 1 Status = 96.05%</b> ALP sites are functional. [73 out of the targeted 76 ALP sites are functional]</p> <p>Overall the project established 73 ALP centres enrolling 1332 OOS girls. The overwhelming demand for ALP is a clear demonstration of how the communities are valuing education especially for girls. The majority [72%] of the girls enrolled in ALP, dropped out of the school between 2014-2017, the effect of drought and migration had an adverse effect on enrolment and retention. Despite the improvements in food security over the past 2 years, the number of girls who dropped out of school is still considerably high.</p> <p>Three of the targeted 76 ALP sites have failed to meet the expected minimum enrolment. Additional sites have been identified and a feasibility assessment will be concluded by end of March.</p>
<b>Output 1.2: Percentage of ALP groups providing life skills training to marginalised girls</b>	Girls are learning relevant life skills that will not only boost their learning outcomes and attendance, but will also enable them to contribute to the local economy once they leave school. This intervention boost learning outcomes, as well as increasing the likelihood of girls transiting into formal schools. Life skills	<p><b>Midline Status =96.05%</b></p> <p>ALP teachers were trained on the ALP curriculum in July 2018 and follow up coaching is ongoing. The findings from the routine monitoring show that all ALP sites are offering life skills lessons as per the curriculum. Life skills is rated by the girls as the most exciting component of the ALP curriculum.</p>


	<p>remain a key component of the program it is vital to constantly monitor the delivery of life skills training.</p> <p>The indicator is still valid; modification is not required.</p>	
<p><b>Output 1.3: Percentage of girls receiving partial grants who remain in school</b></p>	<p>The provision of partial grants to girls enabled girls who are at risk out of school to continue with their education. The project noted increased attendance and retention, almost all partial grants recipient who received grants in 2017, transited to the next grade. The continued support will improve transition rates and learning outcomes for partial grants recipients.</p> <p>The indicator is still relevant.</p>	<p><b>Midline Wave 2- Status: 99.5%</b></p> <p>The project has conducted two rounds of Partial Grant Fidelity of Implementation [FOI]. The first FOI was conducted in May 2018 and a follow up FOI in February 2019. The PG FOI results were used to assess PG processes and to determine the effectiveness of the intervention in improving attendance and retention. The assessment targeted 5 caregivers per school/community. At school level, attendance data was collected through a review of attendance records for each partial grant recipient. Findings from the two rounds of data collection show significant improvements in PG processes, usage of partial grants, contribution of PG in improving attendance and retention. The FOI results show that almost all [99.5%] the partial grants recipients are still in school. The proportion of girls who never missed school was 69.2% in May 2018 and increased to 73.5% in Feb 2019 [+4.3% percentage points from the May 2018 FOI]. In Feb 2019, 85% of caregivers indicated that they received PG on time against 82.8% care givers interviewed in May 2018. In the last disbursement 3.5% faced challenges in accessing PG against 4.1% in May 2018. 83.2% of the caregivers interviewed in 2019 confirmed that they received message on time against 56.8% in May 2018. Other results are as follows;</p>

		<p>- Caregivers who were able to read message [62.76% in February 2019 against 47% in May 2018],</p> <p>-97.3% of caregivers were satisfied or very satisfied with partial grants experience against 71.6% reported in May 2018.</p> <p>-In both assessments the usage of PG for education purposes was at least 99%.</p>
<b>Output 2: Supportive school practices and conditions for marginalised girls</b>		
<b>Output 2.1: Percentage of teachers not using corporal punishment</b>	Addressing corporal punishment will improve conditions for learning this enable girls to attend schools regularly and improve their learning outcomes. The prevalence of corporal punishment should continue to be monitored. The indicator is still valid.	<p><b>Midline Status Wave 1 =68.7%</b></p> <p>Findings from midline evaluation show mixed results on the use of corporal punishment were mixed, though most suggest that corporal punishment declined in intervention schools from baseline to midline. Forcier researchers observed the use of corporal punishment in 66.2 percent of intervention classrooms at the baseline, but just 31.3 percent at the midline. The prevalence of corporal punishment has declined in another measure - the share of students who report that their teacher used physical punishment <i>on any student</i> in the last two weeks, which declined by 13.2 percentage points in intervention schools from baseline to midline.</p>
<b>Output 2.2: Percentage of teachers using the digital learning platform</b>	The digital learning platform is expected to improve the quality of teaching, this will increase student performance and motivation is likely to have a positive effect on attendance and learning.	<p><b>Midline Wave 1 Status = 0</b></p> <p>Prior to the midline the project developed digital content and conducted a pilot for the digital platform in selected schools in SL. The pilot targeted ALP/ in-school girls and their respective teachers. The content developed using interactive methodologies aimed at assessing the girls' understanding and level of participation in the lessons presented. It also focused on gauging the</p>

		pace/timing of the sessions and pronunciation among others.
<b>Output 2.3 Percentage of Girls' Empowerment Forums providing life skills sessions according to the guidance</b>	<p>Life skills – The girls or boys led activities boost their voice and self-confidence, enhancing classroom participation and improved learning outcomes.</p> <p>The indicator is still relevant</p>	<p><b>Midline Wave 1 GEF Status = 31.2%</b>  <b>Midline Wave 1 BEF Status = 66.7%</b></p> <p>This indicator assesses the activities led by girls or boys in their school or communities; these activities are designed to build girl or boys' confidence and participation in the classroom. The data for this indicator was collected through the GEF/BEF fidelity of implementation checklist. The fidelity checklist asked a series of questions to understand the various activities implemented by GEF/BEF; the seven activities assessed include facilitation, debating sessions, competitions, fundraising, sanitation campaigns, community sensitization on girls' education, following up on girls who dropped out of school and participation in other community-related activities.</p> <p>GEF/BEF's who implemented at least 4 out of 7 of the activities, were considered to have met the fidelity of implementation minimum standards.</p> <p>The most common activity for GEF [ 93.8%] and BEF [ 100%] was Sanitisation Campaigns and the least common is debating sessions (GEF [ 37.5%] and BEF[ 22.2%] )</p>
<b>Output 3: Positive shifts on gender and social norms at community and individual girl level</b>		
<b>Output 3:1 Number of women mentors providing support to marginalised girls</b>	<p>Girls who receive mentorship skills from women mentors, will be better equipped to participate in class, breaking traditional norms that restrict girls' voice; to engage in the local economy; and to contribute to their communities in the</p>	<p><b>Midline Wave 1 Status = 236</b></p> <p>The results of the fidelity of implementation show that 75% of the women mentors provided support to the GEFs to undertake various activities which include; facilitating; debating sessions, competitions, fund</p>



	<p>future. This is expected to contribute to improvements in learning and transition outcomes.</p> <p>Indicator still valid.</p>	<p>raising sanitation campaigns, community sensitization on girls' education, following up on girls who dropped out of school and participation in other community-related activities.</p>
<p><b>Output 3:2 Number and percentage of mothers completing literacy courses</b></p>	<p>Mothers in NFE classes acquire essential literacy skills that enable them to support their girls with homework. This will ultimately improve the girl's learning outcome. Skills learnt from the NFE classes will enable them to venture into business, improving their financial capacity to meet the basic education necessities. Girls with adequate basic education necessities are likely to attend school regularly, learn and improve their learning outcomes.</p> <p>Indicator is still valid.</p>	<p><b>Midline Wave 1 Status = 6595</b></p> <p>The project's theory of change assumed that increased parental literacy would contribute to supporting girls and boys to access, stay in school and learn as well as acquiring essential life skills for business/ diversifying livelihoods. A fidelity of implementation assessment was conducted in February 2018 to assess the effectiveness of the NFE program. 122 NFE mothers participated in the assessment. The NFE FOI results show that overall the NFE classes had a significant impact in improving literacy for mothers. Over 97.2% of the NFE mothers regarded NFE as very important or extremely important. Approximately 70.9% of the NFE mothers valued more the ability to read and 25.6% intend to use the knowledge learnt to run or improve their businesses. Sixty three percent of the mothers interviewed indicated that they are now able to assist their children with homework. This is important for improving learning outcomes of marginalised girls in school.</p>
<p><b>Output 3:3 Percentage of active village savings groups in project areas</b></p>	<p>Increased financial capacity of vulnerable households, are expected to contribute to improvements in attendance transition and learning outcomes.</p> <p>Indicator is still valid.</p>	<p><b>Midline Wave 1 Status = 100%</b></p> <p>Based on the preliminary analysis of the 88 VSLA groups' level of functionality and performance, the results show that all the groups established by the project are still functional with a total of 1796 members, cumulative savings amounting to \$56807.17 USD, of which \$ 38,191.95 was cash at hand. Despite the groups having a healthy portfolio, 66% of funds were</p>

		<p>not utilised at the time of the assessment. This is to be expected given the recent establishment of the groups, most of which were established between July to December 2018. In the next phase from January- June 2019, the groups will be trained on the SPM module [Selection, Planning and Management]. The training will enable the groups to start income generation activities. The attached document provide additional information about the level of VSLA functionality and overall portfolio performance.</p> <p> VSLA Functionality and Performance.doc</p>
<p><b>Output 4: Enhanced MOEs' capacity to deliver quality and relevant formal and informal education</b></p>		
<p><b>Output 4:1 Number of Gender Units conducting activities to promote girls' transition and learning</b></p>	<p>Enhancing the capacity of MoEs to develop plans, administer trainings, and provide incentives will contribute to all four intermediate outcomes by sending a strong, positive message about the importance of girls' education from the government, and by giving the government clear and actionable ways to contribute to positive changes in girls' education outcomes.</p> <p>Indicator is still valid.</p>	<p><b>Midline Wave 1 Status = 3</b></p> <p>The project continued to provide Incentives to 13 all-female Gender Focal Person's (GFPs) in the project zones (6 each in SL and PL, 1 in GM). The GFPs worked closely with the project officers to conduct activities earmarked to promote girls' transition. The GFPs were instrumental in the roll out of the ALP component, where they worked closely with the CEC's in mobilisation and enrolment of OOS girls into ALP. Through the reporting period the GFP continued to provide capacity building support to women mentors and school GFP who interact with GEF's on regular basis.</p>

<p><b>Output 4:2 Percentage of REOs/ DEOs engaged in joint monitoring visits to formal schools/ ALP classes to support teachers</b></p>	<p>Regular joint field monitoring visits will improve the quality of project delivery more importantly the quality of teaching. This is expected to lead to improvements in attendance, transition and learning outcomes. Project will take timely adaptations to ensure identified gaps in programming there by improving the quality interventions ultimately contributing to project outputs, intermediate and outcomes.</p> <p>Indicator is still valid.</p>	<p><b>Midline Wave 1 Status = 60%</b></p> <p>MoE Supervisors with the support of REOs and DEOs participated in joint monitoring visits/fidelity of implementation assessments with project staff. The Supervisor across the three project zones visited 89 of the 148 project schools/communities. The joint monitoring visits included an in-depth assessment of the program strengths and identification of areas for improvement. At each school/community the team had in-depth discussions with various stakeholders including care givers for partial grants recipients, mothers participating in NFE and VSLA, CECs members, boys and girls, numeracy, NFE, ALP and head teachers. The information gathered through joint monitoring/FOI assessment inform the realignment of interventions and the bulk of the information was also used to establish the midline status of the output indicators.</p>
---	--	--

List all issues with the means of verification/sources or the frequency of data collection which require changes or additions.

*Table 94: Output indicator issues*

Logframe Output Indicator	Issues with the means of verification/sources and the collection frequency, or the indicator in general?	Changes/additions
<b>Number and Indicator wording</b>	E.g. inappropriate wording, irrelevant sources, or wrong assumptions etc. Was data collection too frequent or too far between? Or no issues?	E.g. change wording, add or remove sources, increase/decrease frequency of data collection; or leave as is.
<b>Output 1: Improved access to post-primary options</b>		
<b>Output 1.1: Percentage of project locations with an alternative learning program for upper primary/ secondary</b>	None	None
<b>Output 1.2: Percentage of ALP groups providing life skills training to marginalised girls</b>	None	None
<b>Output 1.3: Percentage of girls receiving partial grants who remain in school</b>	None	None
<b>Output 2: Supportive school practices and conditions for marginalised girls</b>		
<b>Output 2.1: Percentage of teachers not using corporal punishment</b>	None	None
<b>Output 2.2: Percentage of teachers using the digital learning platform</b>	None	None
<b>Output 2.3 Percentage of Girls' Empowerment Forums providing life skills sessions according to the guidance</b>	None	None
<b>Output 3: Positive shifts on gender and social norms at community and individual girl level</b>		
<b>Output 3:1 Number of women mentors providing support to marginalised girls</b>	None	None
<b>Output 3:2 Number and percentage of mothers completing literacy courses</b>	None	None
<b>Output 3:3 Percentage of active village savings groups in project areas</b>	None	None
<b>Output 4: Enhanced MOEs' capacity to deliver quality and relevant formal and informal education</b>		
<b>Output 4:1 Number of Gender Units conducting activities to</b>	None	None

<b>promote girls' transition and learning</b>		
<b>Output 4:2 Percentage of REOs/ DEOs engaged in joint monitoring visits to formal schools/ ALP classes to support teachers</b>	None	None

## Annex 9: Beneficiaries tables

**Note from CARE:** In light of (i) not having included out-of-school girls in the present assessment; and (ii) having used a smaller and less representative sample due to the inability to access conflict-affected areas, the project has opted for including the same figures used at the baseline in this report. It is not possible to revise the figures at the moment without having access to data on out-of-school girls potentially enrolled/ re-enrolled in formal school (noting that there are multiple monitoring reports indicating that this has occurred) and without having complete data on enrolment due to accessibility issues.

**Table 95: Direct beneficiaries**

Beneficiary type	Total project number	Total number of girls targeted for learning outcomes that the project has reached by Endline	Comments
<b>Direct learning beneficiaries (girls)</b> – 26,290 girls – learning beneficiaries (estimated number of cohort girls attending school or ALP), out of whom 1,814 are girls with disabilities	32,862	[This may equal the total project number in the outcomes spreadsheet and in the column to the left, or may be less if you have a staggered approach]	Overall reach is calculated based on (i) an extrapolation of the enrolment data for 145 schools; (ii) 1332 girls enrolled in ALP and (iii) an estimate of the new intake in Grade 1 (conservatively estimated as equal to the current enrolment, thus avoiding double-counting with OOSG).

*Project note: Calculations*

The number of beneficiaries is calculated as the total at the time of the baseline plus estimates for new enrollees, including 16,689 girls in primary school; 1,912 girls in secondary school; 5,140 out-of-school girls; and 9,121 new entrants. These estimates have not been modified in this evaluation round given the fact that out of school girls were not assessed and it is therefore not possible to obtain data on the numbers benefitting from re-enrolment/ new enrolment.

**Table 96: Other beneficiaries**

Beneficiary type	Number	Comments
<b>Learning beneficiaries (boys)</b> – as above, but specifically counting boys who will get the same exposure and therefore be expected to also achieve learning gains, if applicable.	15,910	Considering 80% of the boys in school as learning beneficiaries.
<b>Broader student beneficiaries (boys)</b> – boys who will benefit from the interventions in a less direct way, and therefore may benefit from aspects such as attitudinal change, etc. but not necessarily achieve improvements in learning outcomes.	30,053	Considering all boys in school, plus new intake in G1.
<b>Broader student beneficiaries (girls)</b> – girls who will benefit from the interventions in a less direct way, and therefore may benefit from aspects such as attitudinal change, etc. but not necessarily achieve improvements in learning outcomes.	27,722	Considering all girls currently enrolled in school, plus new intake in G1.
<b>Teacher beneficiaries</b> – number of teachers who benefit from training or related interventions. If possible /applicable, please disaggregate by gender and type of training, with the comments box used to describe the type of training provided.	621 teachers trained on literacy, numeracy, English and structured remedial classes 158 teachers trained to deliver ASLP	
<b>Broader community beneficiaries (adults)</b> – adults who benefit from broader interventions, such as community messaging /dialogues, community advocacy, economic empowerment interventions, etc.	6595 mothers receiving NFE training 3,180 community members participating in VSLA	

- Tables 3-6 provide different ways of defining and identifying the project's target groups. They each refer to the same total number of girls, but use different definitions and categories. These are girls who can be counted and have regular involvement with project activities.
- The total number of sampled girls in the last row of Tables 3-6 should be the same – these are just different ways of identifying and describing the girls included in the sample.

**Table 97: Target groups - by school**

	Project definition of target group	Number targeted through project interventions	Sample size of target group at Baseline

School Age	(Tick where appropriate)		
Lower primary	Yes - Grade 1-4	19,989	272
Upper primary	Yes - Grade 5-8	5,820	233 (+93 in benchmark)
Lower secondary	Yes - Form 1-2	1,912	12 (benchmark only)
Upper secondary			
<b>Total:</b>			[This number should be the same across Tables 3, 4, 5 & 6]

**Table 98: Target groups - by age**

Age Groups	Project definition of target group (Tick where appropriate)	Number targeted through project interventions	Sample size of target group at Baseline
Aged 6-8 (% aged 6-8)		9,120	This group will benefit from teacher training, improved school management and conditions at the household. However, the sample tracks only girls age 10-19.
Aged 9-11 (% aged 9-11)	√	6,885	241
Aged 12-13 (% aged 12-13)	√	6,885	268
Aged 14-15 (% aged 14-15)	√	5,461	192
Aged 16-17 (%aged 16-17)	√	3,086	109
Aged 18-19 (%aged 18-19)	√	1,425	62
Aged 20+ (% aged 20 and over)			
<b>Total:</b>		<b>32,862</b>	[This number should be the same across Tables 3, 4, 5 & 6]

**Table 99: Target groups - by sub group**

Social Groups	Project definition of target group (Tick where appropriate)	Number targeted through project interventions	Sample size of target group at Baseline
Disabled girls (please disaggregate by disability type)	√	<b>1,814</b>	<b>60</b>
Vision impairment	√	197	5
Hearing impairment	√	230	6

<b>Social Groups</b>	<b>Project definition of target group</b> (Tick where appropriate)	<b>Number targeted through project interventions</b>	<b>Sample size of target group at Baseline</b>
Mobility impairment	√	230	6
Cognitive impairment	√	263	7
Self-care impairment	√	230	6
Communication impairment	√	329	9
Mental health impairment	√	1709	45
Orphaned girls	√	3,615	96
Pastoralist girls	√	3,943	105
Child labourers	√	<b>87<sup>330</sup></b>	<b>13</b>
Poor girls	√	<b>32,862</b>	<b>872</b>
Other (please describe)			
<b>Total:</b>			[This number should be the same across Tables 3, 4, 5 & 6]

**Table 100: Target groups - by school status**

<b>Educational sub-groups</b>	<b>Project definition of target group</b> (Tick where appropriate)	<b>Number targeted through project interventions</b>	<b>Sample size of target group at Baseline</b>
Out-of-school girls: have never attended school	√	1285	92
Out-of-school girls: have attended school, but dropped out	√	3855	275
Girls in-school	√	27722	505
<b>Total:</b>		<b>32,862</b>	[This number should be the same across Tables 3, 4, 5 & 6]

**Table 101: Beneficiaries matrix**

<sup>330</sup> This number does not include girls who support the family business or do unpaid work at home.



Outcomes	Direct beneficiaries			Indirect beneficiaries				
	In-school girls (6-10 grade)	OSG (6-9 years)	OSG (18-25)	In-school boys	HT/Teachers	Parents	SMC/P TA	Local government
Learning	✓			✓	✓	✓		
Transition	✓	✓	✓	✓	✓	✓		
Sustainability	✓	✓	✓		✓	✓	✓	
IO 1: Attendance					✓	✓		
IO 2: Self-esteem and empowerment	✓	✓	✓					
IO3: Parental engagement	✓	✓	✓			✓		
IO4: Quality of teaching	✓				✓	✓	✓	✓
IO5: School management and governance	✓				✓	✓	✓	✓

## Annex 10: MEL Framework

Provide latest, FM-approved version of the MEL Framework as a separate document.

## Annex 11: External Evaluator's Inception Report (where applicable)

Provide latest version of the External Evaluator's Inception Report as a separate document.

## Annex 12: Data collection tools used for Midline

Provide all data collection tools as separate documents.

Provide 1-2 English language transcripts of qualitative sessions.

## Annex 13: Datasets, codebooks and programs

All cleaned and labelled datasets are in Stata format. Replication code is provided in the form of Stata .do files to support the replication of key baseline learning and transition findings, including all outcomes spreadsheet tables. The codebook below provides a summary of key variables for the merged household and learning assessment dataset.

## Codebook

Table 101: Key analysis variables in household survey and learning assessments

Variable	Variable name in dataset	Comments
Girl ID	uniqueid	Unique ID for every girl in the dataset (baseline cohort girls, midline re-contacted girls, midline replacement girls, and ALP girls). Values of uniqueid are shared across waves by the same girl or their replacement (i.e. a girl who is replaced at midline shares uniqueid with her replacement).
Zone	a_5	Coded by enumerators prior to household survey start.
Region	Pre_5	Coded by enumerators prior to household survey start.
District	Pre_6	Coded by enumerators prior to household survey start.
Village	Pre_7	Coded by enumerators prior to household survey start.
School	a_7	Coded by enumerators prior to household survey start. Uniquely identifies schools across rounds – codes are shared by same school in baseline and midline.
Age	age	Both girls and caregivers were asked the age of the girl. Where responses from girls and caregivers contradicted one another, data from the baseline was used to triangulate the correct age (i.e. choosing the girl or caregiver response that produces a logical progression in age from baseline). In cases where this process was not dispositive, caregivers were re-contacted by researchers to verify the girl's age.
Grade	grade	Grades range from 1 to 8; grade value equal to 0 indicates that girl is not enrolled.  Both girls and caregivers were asked the girl's grade. Where responses from girls and caregivers contradicted one another, data from the baseline and the head teacher (school) survey was used to triangulate the correct grade (i.e. choosing the girl or caregiver response that produces a logical progression in grade from baseline; checking grade provided by girl and caregiver against school records obtained from head teacher). In cases where this process was not dispositive, caregivers were re-contacted by researchers to verify the girl's grade.
Enrolment status	enrol	Both girls and caregivers were asked about enrolment status. Where responses from girls and caregivers contradicted one another, data from the head teacher (school) survey, specifically school records, were used to adjudicate. In cases where this process was not dispositive, caregivers were re-contacted by researchers to verify the girl's enrolment status.
Intervention and comparison groups	intervention	Intervention and comparison designation is coded here.
Evaluation round	round	Identifies the round (baseline versus midline)

Respondent type	resp_type	Identifies the type of girl and observation. 1 = Cohort girls, baseline data 2 = Cohort girls, successfully re-contacted, midline data 3 = Replacement girls, where cohort girl could not be re-contacted, midline data 4 = ALP (Alternative Learning Programme) girls, midline data  Also used to identify replacement versus re-contacted girls at midline (i.e. to calculate re-contact rates).
School panel sample	school_sample	Identifies (=1) schools that appear in both baseline and midline data; does not identify specifically <i>girls</i> who appear in both rounds
Main girl sample	sample	Identifies (=1) girls in the main cohort sample from baseline to midline. Includes all girls who were either re-contacted or replaced at midline. 807 girls from baseline; 648 girls re-contacted successfully at midline; 159 girls replaced at midline. Total sample = 1,614 girl-observations.
"True panel" sample	panel	Identifies (=1) girls who are part of the "true panel" of cohort girls, i.e. those girls who appear in both baseline and midline samples and are <i>the same</i> girl. No replacement girls from midline appear. 648 cohort girls at baseline and 648 re-contacted cohort girls at midline.
<b>Learning Outcomes</b>		
Somali Literacy Score (percent)	somlit_perc	Calculated from learning assessment
Numeracy Score (percent)	numeracy_perc	Originated from learning assessment data.
Baseline Combined Literacy Score (percent)	bl_literacy_perc	Originated from baseline learning assessment; combines Somali and English literacy scores from baseline. Not provided for midline, because English literacy was not assessed.
English Literacy Score (percent) at baseline	bl_englit_perc	Originated from baseline learning assessment; not provided for midline, because English literacy was not assessed.
Numeracy subtask 1 (ML)	num_sb1	Midline only
Numeracy subtask 2 (ML)	num_sb2	Midline only
Numeracy subtask 3 (ML)	num_sb3	Midline only
Numeracy subtask 4 (ML)	num_sb4	Midline only
Numeracy subtask 5 (ML)	num_sb5	Midline only
Numeracy subtask 6 (ML)	num_sb6	Midline only
Numeracy subtask 7 (ML)	num_sb7	Midline only
Numeracy subtask 8 (ML)	num_sb8	Midline only
Numeracy subtask 9 (ML)	num_sb9	Midline only
Numeracy subtask 10 (ML)	num_sb10	Midline only
Numeracy subtask 11 (ML)	num_sb11	Midline only
Numeracy subtask 1 (BL)	bl_pnum_missing_num_total	Baseline only
Numeracy subtask 2 (BL)	bl_pnum_add_total	Baseline only
Numeracy subtask 3 (BL)	bl_pnum_sub_total	Baseline only

Numeracy subtask 4 (BL)	bl_pcnun_add2_total	Baseline only
Numeracy subtask 5 (BL)	bl_pcnun_sub2_total	Baseline only
Numeracy subtask 6 (BL)	bl_pcnun_wprob_total	Baseline only
Numeracy subtask 7 (BL)	bl_pcnun_mult_total	Baseline only
Numeracy subtask 8 (BL)	bl_pcnun_mult2_total	Baseline only
Numeracy subtask 9 (BL)	bl_pcnun_div_total	Baseline only
Numeracy subtask 10 (BL)	bl_pcnun_div2_total	Baseline only
Numeracy subtask 11 (BL)	bl_pcnun_wprob2_total	Baseline only
Somali literacy subtask 1 (ML)	lit_sb1	Midline only
Somali literacy subtask 2 (ML)	lit_sb2	Midline only
Somali literacy subtask 3 (ML)	lit_sb3	Midline only
Somali literacy subtask 4 (ML)	lit_sb4	Midline only
Somali literacy subtask 5 (ML)	lit_sb5	Midline only
Somali literacy subtask 6 (ML)	lit_sb6	Midline only
Somali literacy subtask 7 (ML)	lit_sb7	Midline only
Somali literacy subtask 8 (ML)	lit_sb8	Midline only
Somali literacy subtask 1 (BL)	bl_prop_wpm_score_som1	Baseline only
Somali literacy subtask 2 (BL)	bl_score_pct_som2	Baseline only
Somali literacy subtask 3 (BL)	bl_score_pct_som3	Baseline only
Somali literacy subtask 4 (BL)	bl_prop_wpm_score_som4	Baseline only
Somali literacy subtask 5 (BL)	bl_score_pct_som5	Baseline only
Somali literacy subtask 6 (BL)	bl_score_pct_som6	Baseline only
Somali literacy subtask 7 (BL)	bl_score_pct_som7	Baseline only
Somali literacy subtask 8 (BL)	bl_score_pct_som8	Baseline only
<b>Transition Outcomes</b>		
Transition pathway	transition_path	<p>Derived from enrol (enrolment status) and grade variables from baseline, midline, and retrospective at midline, where appropriate.</p> <p>Note: to facilitate transparency and review of its construction, this variable is coded in the replication .do files and is not included in the datasets themselves. This allows replicators to see the exact coding decisions made by reviewing the .do file.</p>
Binary transition indicator	transition	<p>Derived from transition_path, this is a binary variable indicating transition success (1) or failure (0). Note that it is also coded in the replication .do files included in this annex.</p>
Simplified set of transition pathways	transition_simp	<p>A simplified version of transition_path, which consolidates some pathways to facilitate analysis where few girls entered some of the pathways. Combines alternative education, vocational training, and employment into a single category. Coded in replication .do files.</p>

## Key Regression Results

### Learning outcomes for cohort girls

#### Difference-in-differences in literacy for in-school cohort girls (with replacements)

```

Linear regression              Number of obs   =    1,388
                              F(3, 62)       =    38.01
                              Prob > F              =    0.0000
                              R-squared             =    0.0960
                              Root MSE          =    28.015
  
```

(Std. Err. adjusted for 63 clusters in a\_7)

lit	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
round Midline	18.20914	2.323558	7.84	0.000	13.56441	22.85387
treatment Treatment	4.287529	3.644016	1.18	0.244	-2.996754	11.57181
round#treatment Midline#Treatment	-.735723	3.348435	-0.22	0.827	-7.429148	5.957702
_cons	41.26705	2.977519	13.86	0.000	35.31508	47.21903

#### Difference-in-differences in literacy for in-school cohort girls (panel)

```

Linear regression              Number of obs   =    1,128
                              F(3, 62)       =    37.96
                              Prob > F              =    0.0000
                              R-squared             =    0.0971
                              Root MSE          =    28.219
  
```

(Std. Err. adjusted for 63 clusters in a\_7)

lit	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
round Midline	18.43849	2.296706	8.03	0.000	13.84744	23.02953
treatment Treatment	4.512937	3.755577	1.20	0.234	-2.994354	12.02023
round#treatment Midline#Treatment	-.7995581	3.41106	-0.23	0.815	-7.61817	6.019054
_cons	41.93502	2.977319	14.08	0.000	35.98344	47.8866

### Difference-in-differences in numeracy for in-school cohort girls (with replacements)

Linear regression

Number of obs	=	1,388
F(3, 62)	=	10.78
Prob > F	=	0.0000
R-squared	=	0.0605
Root MSE	=	25.248

(Std. Err. adjusted for 63 clusters in a\_7)

num	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
round Midline	12.74034	2.907893	4.38	0.000	6.927547	18.55314
treatment Treatment	2.579764	4.418644	0.58	0.561	-6.25298	11.41251
round#treatment Midline#Treatment	-.3161004	4.495545	-0.07	0.944	-9.302567	8.670366
_cons	38.74092	3.200341	12.11	0.000	32.34353	45.13831

### Difference-in-differences in numeracy for in-school cohort girls (panel)

Linear regression

Number of obs	=	1,128
F(3, 62)	=	12.37
Prob > F	=	0.0000
R-squared	=	0.0651
Root MSE	=	25.141

(Std. Err. adjusted for 63 clusters in a\_7)

num	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
round Midline	13.12215	2.760988	4.75	0.000	7.603012	18.64128
treatment Treatment	3.23845	4.023176	0.80	0.424	-4.803763	11.28066
round#treatment Midline#Treatment	-.4157399	4.414897	-0.09	0.925	-9.240993	8.409513
_cons	37.72727	2.66936	14.13	0.000	32.3913	43.06325

## Comparison of boys' and girls' learning outcomes

### Literacy outcomes for in-school learners, by gender

Linear regression

Number of obs	=	938
F(1, 70)	=	1.77
Prob > F	=	0.1875
R-squared	=	0.0025
Root MSE	=	29.246

(Std. Err. adjusted for 71 clusters in a\_7)

lit	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
gender						
Boys	3.671295	2.758024	1.33	0.187	-1.82941	9.172
_cons	59.52092	2.049249	29.05	0.000	55.43383	63.60802

### Numeracy outcomes for in-school learners, by gender

Linear regression

Number of obs	=	938
F(1, 70)	=	14.78
Prob > F	=	0.0003
R-squared	=	0.0140
Root MSE	=	25.311

(Std. Err. adjusted for 71 clusters in a\_7)

num	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
gender						
Boys	7.642572	1.987893	3.84	0.000	3.677844	11.6073
_cons	51.59902	1.915028	26.94	0.000	47.77961	55.41842

### Test of specific intervention-variables

#### DID in numeracy outcomes for girls in classes with teachers trained on numeracy-boost (with replacements)

Linear regression

Number of obs	=	762
F(3, 31)	=	5.10
Prob > F	=	0.0055
R-squared	=	0.0592
Root MSE	=	24.914

(Std. Err. adjusted for 32 clusters in a\_7)

num	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
round						
Midline	11.67088	3.000905	3.89	0.000	5.550489	17.79126
1.teacher_training	-.2546678	2.995631	-0.09	0.933	-6.364297	5.854962
round#teacher_training						
Midline#1	1.739593	3.737156	0.47	0.645	-5.882387	9.361573
_cons	41.43098	3.604274	11.49	0.000	34.08001	48.78194

#### DID in numeracy outcomes for girls in classes with teachers trained on numeracy-boost (panel only)

Linear regression

Number of obs	=	610
F(3, 31)	=	5.96
Prob > F	=	0.0025
R-squared	=	0.0616
Root MSE	=	24.939

(Std. Err. adjusted for 32 clusters in a\_7)

num	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
round						
Midline	13.27479	3.146757	4.22	0.000	6.85694	19.69265
1.teacher_training	1.349042	2.879496	0.47	0.643	-4.52373	7.221813
round#teacher_training						
Midline#1	-1.343856	3.756257	-0.36	0.723	-9.004793	6.317081
_cons	40.39514	3.47883	11.61	0.000	33.30003	47.49026





**At midline, numeracy outcomes for girls who reported participating in GEF as compared with girls who reported not participating (panel only)**

Linear regression

Number of obs	=	561
F(1, 62)	=	14.96
Prob > F	=	0.0003
R-squared	=	0.0351
Root MSE	=	24.869

(Std. Err. adjusted for 63 clusters in a\_7)

num	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
gef	11.96182	3.09314	3.87	0.000	5.778722	18.14492
_cons	50.13576	1.966568	25.49	0.000	46.20465	54.06687

**Literacy outcomes for girls who reported participating in GEF as compared with girls who reported not participating, baseline versus midline (with replacements)**

Linear regression

Number of obs	=	1,353
F(3, 62)	=	52.51
Prob > F	=	0.0000
R-squared	=	0.1131
Root MSE	=	27.81

(Std. Err. adjusted for 63 clusters in a\_7)

lit	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
round#gef						
Baseline#1	13.81491	7.396982	1.87	0.067	-.9714427	28.60127
Midline#0	16.62887	1.805653	9.21	0.000	13.01943	20.23832
Midline#1	28.93588	2.724749	10.62	0.000	23.48918	34.38257
_cons	42.60185	1.855837	22.96	0.000	38.89209	46.31162



**Numeracy outcomes for girls who reported participating in GEF as compared with girls who reported not participating, baseline versus midline (panel only)**

Linear regression

Number of obs	=	1,100
F(3, 62)	=	28.61
Prob > F	=	0.0000
R-squared	=	0.0932
Root MSE	=	24.708

(Std. Err. adjusted for 63 clusters in a\_7)

num	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
round#gef						
Baseline#1	18.59788	5.647936	3.29	0.002	7.30782	29.88795
Midline#0	11.65031	2.190617	5.32	0.000	7.271329	16.02929
Midline#1	23.61213	3.325931	7.10	0.000	16.96369	30.26057
_cons	38.48545	1.980055	19.44	0.000	34.52738	42.44352

**Supplemental Learning Gap Tables**  
**Girls' intervention numeracy**

Categories	Subtask 1	Subtask 2	Subtask 3	Subtask 4	Subtask 5	Subtask 6	Subtask 7	Subtask 8	Subtask 9	Subtask 10	Subtask 11
	Missing Number	Addition (Level 1)	Subtraction (Level 1)	Addition (Level 2)	Subtraction (Level 2)	Word problems (add/subtract)	Multiplication (Level 1)	Multiplication (Level 2)	Division (Level 1)	Division (Level 2)	Word problems (mult/div)
<b>Non-learner 0%</b>	2.3 (-1.1)	2.7 (-3.1)	6 (-18.5)	15.3 (-18.9)	23.6 (-24.4)	11.5 (-21.9)	18.1 (-28.3)	69.6 (-6.9)	49.1 (-22)	75.9 (-9.2)	58.9 (-11)
<b>Emergent learner 1%-40%</b>	37.1 (7.1)	2.8 (-1.4)	3.9 (-3.5)	15.4 (-6.1)	10.6 (-3)	5.2 (-1.9)	15.3 (0.8)	12.4 (3.9)	15 (4)	11.8 (5.6)	0 (0)
<b>Established learner 41%-80%</b>	37.8 (9.4)	17.8 (-1)	12.1 (-3)	34.1 (12)	33.9 (12.5)	33.5 (2.5)	41.8 (23)	9.5 (4.7)	16.2 (9.8)	3.9 (0.8)	14 (2.3)
<b>Proficient learner 81%-100%</b>	22.7 (-15.4)	76.7 (5.4)	78 (25)	35.2 (13)	31.9 (15)	49.8 (21.3)	24.7 (4.5)	8.4 (-1.7)	19.8 (8.1)	8.5 (2.8)	27.2 (8.7)
	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

**Girls' comparison numeracy**

<b>Categories</b>	<b>Subtask 1</b> Missing Number	<b>Subtask 2</b> Addition (Level 1)	<b>Subtask 3</b> Subtraction (Level 1)	<b>Subtask 4</b> Addition (Level 2)	<b>Subtask 5</b> Subtraction (Level 2)	<b>Subtask 6</b> Word problems (add/subtract)	<b>Subtask 7</b> Multiplication (Level 1)	<b>Subtask 8</b> Multiplication (Level 2)	<b>Subtask 9</b> Division (Level 1)	<b>Subtask 10</b> Division (Level 2)	<b>Subtask 11</b> Word problems (mult/div)
<b>Non-learner 0%</b>	3.3 (-2.2)	3 (-8.9)	8.1 (-15.3)	18.7 (-19.7)	28.4 (-23.2)	12.6 (-24.7)	22.1 (-28.1)	74.3 (-6.3)	50.1 (-21.8)	81.2 (-5)	57.6 (-15.1)
<b>Emergent learner 1%-40%</b>	37.2 (-2.3)	3.4 (-3.3)	6.8 (-4.2)	21 (-0.4)	12.6 (0.2)	12.2 (4.7)	17.9 (4.3)	9.6 (0.5)	17.9 (6)	7.1 (1.4)	0 (0)
<b>Established learner 41%-80%</b>	41.4 (13)	24.8 (7.5)	16.1 (3.9)	28 (9.6)	30.4 (11.2)	27.8 (1.2)	35.5 (16.6)	11.4 (6)	13.6 (5.5)	6.8 (2.2)	14.5 (2.4)
<b>Proficient learner 81%-100%</b>	18.2 (-8.5)	68.7 (4.6)	69.1 (15.6)	32.2 (10.5)	28.6 (11.9)	47.4 (18.7)	24.5 (7.2)	4.7 (-0.3)	18.4 (10.2)	4.9 (1.4)	27.9 (12.7)
	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

**Girls' DID numeracy by proficiency level**

<b>Categories</b>	<b>Subtask 1</b> Missing Number	<b>Subtask 2</b> Addition (Level 1)	<b>Subtask 3</b> Subtraction (Level 1)	<b>Subtask 4</b> Addition (Level 2)	<b>Subtask 5</b> Subtraction (Level 2)	<b>Subtask 6</b> Word problems (add/subtract)	<b>Subtask 7</b> Multiplication (Level 1)	<b>Subtask 8</b> Multiplication (Level 2)	<b>Subtask 9</b> Division (Level 1)	<b>Subtask 10</b> Division (Level 2)	<b>Subtask 11</b> Word problems (mult/div)
<b>Non-learner 0%</b>	1.1	5.8	-3.2	0.8	-1.2	2.8	-0.2	-0.6	-0.2	-4.2	4.1
<b>Emergent learner 1%-40%</b>	9.4	1.9	0.7	-5.7	-3.2	-6.6	-3.5	3.3	-2.0	4.2	0.0
<b>Established learner 41%-80%</b>	-3.7	-8.5	-6.9	2.4	1.3	1.3	6.3	-1.3	4.3	-1.4	-0.1
<b>Proficient learner 81%-100%</b>	-6.9	0.8	9.4	2.5	3.1	2.5	-2.7	-1.4	-2.1	1.4	-4.0

### Girls' intervention literacy

Categories	Somali ST1	Somali ST2	Somali ST3	Somali ST4	Somali ST5	Somali ST6	Somali ST7	Somali ST8
	Reading Words	Reading Comp (easy)	Reading Comp (medium)	Reading Fluency	Reading Comp (difficult)	Writing (fill blank)	Writing (negative form)	Writing (future tense)
<b>Non-learner 0%</b>	2.8 (-11.1)	9.8 (-5.9)	13.6 (-10.2)	4.9 (-20.4)	25.5 (-12.6)	28.2 (-17.7)	39.8 (-22.3)	46.2 (-21.5)
<b>Emergent learner 1%-40%</b>	5.2 (-41.1)	6.6 (4.4)	10.8 (3.8)	18 (-5.8)	17.2 (2.7)	17.6 (4)	5.2 (0)	7 (3.7)
<b>Established learner 41%-80%</b>	15.8 (-22.9)	40.1 (12.5)	39 (10.3)	36.6 (3.3)	42.7 (14.8)	18.3 (0.2)	7.9 (-3.3)	6.3 (-1.7)
<b>Proficient learner 81%-100%</b>	76.2 (75.1)	43.4 (-11)	36.6 (-3.9)	40.5 (22.9)	14.6 (-4.9)	36 (13.5)	47.1 (25.6)	40.6 (19.5)
	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

### Girls' comparison literacy

Categories	Somali ST1	Somali ST2	Somali ST3	Somali ST4	Somali ST5	Somali ST6	Somali ST7	Somali ST8
	Reading Words	Reading Comp (easy)	Reading Comp (medium)	Reading Fluency	Reading Comp (difficult)	Writing (fill blank)	Writing (negative form)	Writing (future tense)
<b>Non-learner 0%</b>	4.2 (-11.6)	12.8 (-8.6)	19.2 (-10.4)	6.2 (-21.7)	25.2 (-19.6)	27.6 (-20)	40.7 (-20.2)	46.6 (-22.2)
<b>Emergent learner 1%-40%</b>	6.6 (-42.9)	6.8 (3.4)	10.6 (0.4)	16.2 (-8.4)	18.9 (8.4)	17 (3.3)	7.5 (-0.4)	7.4 (3.2)
<b>Established learner 41%-80%</b>	16.4 (-17.8)	34.9 (9.9)	40.2 (10.4)	42.9 (13.5)	38.7 (7.7)	21.5 (7.9)	9.3 (0.8)	10.4 (3.2)
<b>Proficient learner 81%-100%</b>	72.8 (72.4)	45.5 (-4.6)	29.9 (-0.4)	34.7 (16.6)	17.2 (3.6)	33.8 (8.8)	42.6 (19.9)	35.6 (15.8)
	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0



### Girls' DID literacy by proficiency level

Categories	Somali ST1 Reading Words	Somali ST2 Reading Comp (easy)	Somali ST3 Reading Comp (medium)	Somali ST4 Reading Fluency	Somali ST5 Reading Comp (difficult)	Somali ST6 Writing (fill blank)	Somali ST7 Writing (negative form)	Somali ST8 Writing (future tense)
<b>Non-learner 0%</b>	0.6	2.8	0.2	1.2	7.0	2.3	-2.1	0.6
<b>Emergent learner 1%-40%</b>	1.8	1.0	3.4	2.7	-5.7	0.7	0.5	0.6
<b>Established learner 41%-80%</b>	-5.1	2.6	-0.1	-10.2	7.1	-7.7	-4.1	-4.9
<b>Proficient learner 81%-100%</b>	2.7	-6.3	-3.5	6.3	-8.4	4.7	5.7	3.7

## Boys' intervention numeracy

Categories	Subtask 1	Subtask 2	Subtask 3	Subtask 4	Subtask 5	Subtask 6	Subtask 7	Subtask 8	Subtask 9	Subtask 10	Subtask 11
	Missing Number	Addition (Level 1)	Subtraction (Level 1)	Addition (Level 2)	Subtraction (Level 2)	Word problems (add/subtract)	Multiplication (Level 1)	Multiplication (Level 2)	Division (Level 1)	Division (Level 2)	Word problems (mult/div)
<b>Non-learner 0%</b>	6.6 (-1.8)	7.5 (-3.2)	10.4 (-22.3)	21.7 (-17.1)	26.4 (-23.1)	17.9 (-18.5)	32.1 (-20.7)	68.9 (-6.4)	53.8 (-16.3)	75.5 (-7.2)	58.5 (-9.7)
<b>Emergent learner 1%-40%</b>	29.2 (1.7)	2.8 (-1.4)	3.8 (-1.4)	17 (0.2)	9.4 (-5.5)	3.8 (-3.7)	5.7 (-5.1)	11.3 (1.5)	11.3 (1.5)	8.5 (1)	0 (0)
<b>Established learner 41%-80%</b>	34.9 (8.7)	19.8 (2.1)	11.3 (0.6)	23.6 (7.2)	26.4 (8.7)	30.2 (6.4)	33 (18.5)	7.5 (3.3)	13.2 (2.9)	7.5 (5.2)	14.2 (2)
<b>Proficient learner 81%-100%</b>	29.2 (-8.6)	69.8 (2.5)	74.5 (23.1)	37.7 (9.7)	37.7 (20)	48.1 (15.9)	29.2 (7.3)	12.3 (1.5)	21.7 (11.9)	8.5 (1)	27.4 (7.7)
	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

### Boys' comparison numeracy

Categories	Subtask 1 Missing Number	Subtask 2 Addition (Level 1)	Subtask 3 Subtraction (Level 1)	Subtask 4 Addition (Level 2)	Subtask 5 Subtraction (Level 2)	Subtask 6 Word problems (add/subtract)	Subtask 7 Multiplication (Level 1)	Subtask 8 Multiplication (Level 2)	Subtask 9 Division (Level 1)	Subtask 10 Division (Level 2)	Subtask 11 Word problems (mult/div)
<b>Non-learner 0%</b>	3.8 (-1.9)	5.1 (-3.4)	5.1 (-14.7)	12.7 (-17.1)	20.3 (-23.1)	6.3 (-22.4)	22.8 (-25.8)	62 (-10.1)	43 (-23)	67.1 (-15.5)	45.6 (-29)
<b>Emergent learner 1%-40%</b>	30.4 (4)	0 (-9)	0 (-4.7)	15.2 (2)	11.4 (0.1)	6.3 (-1.7)	10.1 (-1.2)	20.3 (5.2)	11.4 (-1.8)	13.9 (4.5)	0 (0)
<b>Established learner 41%-80%</b>	34.2 (4.9)	17.7 (4.5)	8.9 (-7.6)	21.5 (-4.4)	35.4 (10.4)	41.8 (8.8)	31.6 (12.3)	12.7 (6.5)	20.3 (8.5)	16.5 (11.3)	21.5 (10.2)
<b>Proficient learner 81%-100%</b>	31.6 (-7)	77.2 (7.9)	86.1 (27.1)	50.6 (19.5)	32.9 (12.6)	45.6 (15.4)	35.4 (14.7)	5.1 (-1.5)	25.3 (16.4)	2.5 (-0.3)	32.9 (18.8)
	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

**Boys' DID numeracy by proficiency level**

<b>Categories</b>	<b>Subtask 1</b> Missing Number	<b>Subtask 2</b> Addition (Level 1)	<b>Subtask 3</b> Subtraction (Level 1)	<b>Subtask 4</b> Addition (Level 2)	<b>Subtask 5</b> Subtraction (Level 2)	<b>Subtask 6</b> Word problems (add/subtract)	<b>Subtask 7</b> Multiplication (Level 1)	<b>Subtask 8</b> Multiplication (Level 2)	<b>Subtask 9</b> Division (Level 1)	<b>Subtask 10</b> Division (Level 2)	<b>Subtask 11</b> Word problems (mult/div)
<b>Non-learner 0%</b>	0.1	0.2	-7.6	0.0	0.0	3.9	5.1	3.8	6.7	8.2	19.2
<b>Emergent learner 1%-40%</b>	-2.3	7.6	3.4	-1.8	-5.6	-2.0	-3.9	-3.7	3.3	-3.5	0.0
<b>Established learner 41%-80%</b>	3.8	-2.5	8.2	11.7	-1.8	-2.4	6.2	-3.2	-5.5	-6.1	-8.2
<b>Proficient learner 81%-100%</b>	-1.6	-5.4	-4.0	-9.8	7.4	0.5	-7.4	3.1	-4.5	1.3	-11.0

### Boys' intervention literacy

Categories	Somali ST1 Reading Words	Somali ST2 Reading Comp (easy)	Somali ST3 Reading Comp (medium)	Somali ST4 Reading Fluency	Somali ST5 Reading Comp (difficult)	Somali ST6 Writing (fill blank)	Somali ST7 Writing (negative form)	Somali ST8 Writing (future tense)
<b>Non-learner 0%</b>	6.6 (-13)	12.3 (-11.1)	20.8 (-9.2)	10.4 (-17.7)	37.7 (-12.3)	35.8 (-10.9)	40.6 (-19.7)	50 (-13.6)
<b>Emergent learner 1%-40%</b>	4.7 (-34.5)	7.5 (2.9)	10.4 (3.4)	20.8 (-2.6)	23.6 (12.4)	15.1 (3.9)	5.7 (0.5)	6.6 (1)
<b>Established learner 41%-80%</b>	21.7 (-16.6)	34.9 (12.9)	30.2 (-0.2)	34 (2.7)	20.8 (-5.9)	14.2 (-0.3)	9.4 (0.6)	8.5 (-1.3)
<b>Proficient learner 81%-100%</b>	67 (64.2)	45.3 (-4.7)	38.7 (6)	34.9 (17.6)	17.9 (5.8)	34.9 (7.3)	44.3 (18.6)	34.9 (13.9)
	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

### Boys' comparison literacy

Categories	Somali ST1 Reading Words	Somali ST2 Reading Comp (easy)	Somali ST3 Reading Comp (medium)	Somali ST4 Reading Fluency	Somali ST5 Reading Comp (difficult)	Somali ST6 Writing (fill blank)	Somali ST7 Writing (negative form)	Somali ST8 Writing (future tense)
<b>Non-learner 0%</b>	3.8 (-12.2)	7.6 (-12.2)	13.9 (-14.8)	10.1 (-18.6)	22.8 (-24.4)	26.6 (-18.2)	35.4 (-19.7)	46.8 (-14)
<b>Emergent learner 1%-40%</b>	2.5 (-41.8)	7.6 (3.8)	8.9 (-2)	11.4 (-11.2)	15.2 (8.6)	13.9 (0.7)	2.5 (-4.5)	3.8 (-4.7)
<b>Established learner 41%-80%</b>	15.2 (-24)	29.1 (3.2)	45.6 (21)	46.8 (12.9)	41.8 (12.1)	24.1 (12.3)	15.2 (6.7)	12.7 (4.2)
<b>Proficient learner 81%-100%</b>	78.5 (78)	55.7 (5.2)	31.6 (-4.2)	31.6 (17)	20.3 (3.7)	35.4 (5.3)	46.8 (17.6)	36.7 (14.5)
	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

## Boys' DID literacy by proficiency level

Categories	Somali ST1 Reading Words	Somali ST2 Reading Comp (easy)	Somali ST3 Reading Comp (medium)	Somali ST4 Reading Fluency	Somali ST5 Reading Comp (difficult)	Somali ST6 Writing (fill blank)	Somali ST7 Writing (negative form)	Somali ST8 Writing (future tense)
<b>Non-learner 0%</b>	-0.8	1.1	5.7	1.0	12.1	7.3	0.0	0.5
<b>Emergent learner 1%-40%</b>	7.3	-0.9	5.4	8.6	3.8	3.2	5.1	5.7
<b>Established learner 41%-80%</b>	7.3	9.8	-21.2	-10.2	-17.9	-12.6	-6.1	-5.5
<b>Proficient learner 81%-100%</b>	-13.8	-9.9	10.2	0.6	2.0	2.1	1.0	-0.7

## Annex 14: Learning test pilot and calibration

### Assessment Design

Prior to the start of midline data collection, CARE's Monitoring and Evaluation team designed the midline learning assessments. While the baseline evaluation included four distinct assessment modules (numeracy, English literacy, Somali literacy, and financial literacy), only two modules were designed for use in this midline – numeracy and Somali literacy. English literacy was excluded from the midline, as noted elsewhere in this report, because interventions targeting English literacy had not begun at the time of the midline, or were sufficiently new that they could not reasonably be expected to have impacted English literacy scores already. The tests were designed under guidance from the GEC FM for the development of SeGRA and SeGMA, mirroring the structure of the Early Grade Reading and Maths Assessments (EGRA and EGMA).

At the midline, learning assessments were developed to ensure assessments of equivalent difficulty to the baseline, to facilitate comparisons over time. The primary difference between the baseline and midline assessments is the exclusion of English literacy from the midline assessment. Within the numeracy and Somali literacy assessments, there have been no changes in terms of number of subtasks, number of items within a given subtask, or grading of subtasks. All adjustments to the midline assessments were minor, consisting of changing the order of words, replacing names in stories, adjusting comprehension questions slightly, and using different numbers in the numeracy assessment, but maintaining the same

number of digits (and therefore equivalent levels of difficulty). The assessments test the following general skills (full-text versions of the assessments are provided in Annex 12):

### *Numeracy*

- Subtask 1: Missing Numbers
- Subtask 2: Addition (level 1)
- Subtask 3: Subtraction (level 1)
- Subtask 4: Addition (level 2)
- Subtask 5: Subtraction (level 2)
- Subtask 6: Addition and subtraction word problems
- Subtask 7: Multiplication (level 1)
- Subtask 8: Multiplication (level 2)
- Subtask 9: Division (level 1)
- Subtask 10: Division (level 2)
- Subtask 11: Multiplication and division word problems

### *Somali Literacy*

- Subtask 1: Reading Fluency – High-Frequency Words
- Subtask 2: Reading comprehension (easy)
- Subtask 3: Reading comprehension (medium)
- Subtask 4: Reading fluency (story reading)
- Subtask 5: Reading comprehension (hard)
- Subtask 6: Writing (fill in missing words)
- Subtask 7: Writing (convert sentence to negative form)
- Subtask 8: Writing (convert sentence to future tense)

Each subtask comprised a set of individual items, ranging from one to ten per subtask. There was no pilot of the midline assessment. Given the efforts made to maintain the core structure of the assessment and ensure equivalent difficulty, a pilot and elaborate calibration process were not seen as necessary – as we discuss below, this view is borne out by the equivalence testing conducted during the midline, which showed that baseline and midline assessments were of similar difficulty.

The scoring methodology ensured that each subtask was weighted equally in the final aggregate score. Specifically, each subtask was scored as the percentage of items correct out of the total number of items (hence ranging from 0 to 100). In keeping with FM guidance, the reading tasks that involved a word-per-minute (WPM) score were censored at a cap of 100 WPM, with individuals who scored above 100 WPM being assigned a score of 100 WPM. The result is that all subtasks were individually standardized to range from 0 to 100. The total score for the numeracy and literacy assessments was then generated by taking the average of the subtask scores for that assessment (with each subtask being given equal weight), presenting the total percentage score based on the averaged subtasks, ranging between 0 and 100. The financial literacy exam is an exception to this scoring procedure, as the first and second sections are scored and presented separately, each having a score ranging between 0 and 100. This procedure ensured that each subtask (and the associated skills) made an equal contribution to the final score for a given assessment, and that the final scores for each assessment have a comparable range from 0 to 100. For further details on assessment scoring and piloting, please see Annex 9.

With regard to target-setting, all evaluation targets for the midline 2 evaluation were set on the basis of the outcomes worksheet.

### Assessment Calibration

No pilot test of the midline learning assessments were conducted prior to the start of fieldwork. Therefore, the discussion in this section focuses on the actual learning assessment implemented at the midline. Calibration of the midline and baseline assessments' difficulty levels occurred during midline data collection. A random subsample of girls was selected to complete both the baseline and midline assessment for a given topic. By having the same girl take both assessments, we can compare their relative performance and draw conclusions about the relative difficulty of the tests.

In total, 42 girls completed both literacy assessments and 38 girls completed both numeracy assessments.<sup>331</sup> The girls participating in the calibration (hereafter "calibration girls") performed similarly to the broader sample of cohort girls in terms of literacy, but not numeracy: their mean score on the Somali literacy midline was 63.5 percent, compared to 60.2 percent in the full sample; on numeracy, their mean score was 59.3 percent, compared to 51.8 percent in full midline sample. Grade representation was balanced, which may help account for the calibration girls' superior performance in numeracy, as girls in older grades were slightly overrepresented among calibration girls, relative to the broader midline sample.

More importantly, the results from calibration girls show the baseline and midline assessments were approximately equivalent in difficulty, as discussed in Section 4.1 of the report. The mean scores on the assessments, among calibration girls, were:

- Baseline numeracy: 59.7% (n = 38)
- Midline numeracy: 59.3% (n = 38)
- Baseline Somali literacy: 62.5% (n = 42)
- Midline Somali literacy: 63.5% (n = 42)

The differences between baseline and midline scores were not statistically significant at any conventional level. A two-sided t-test produced p-values of 0.82 and 0.70 for numeracy and literacy, respectively.

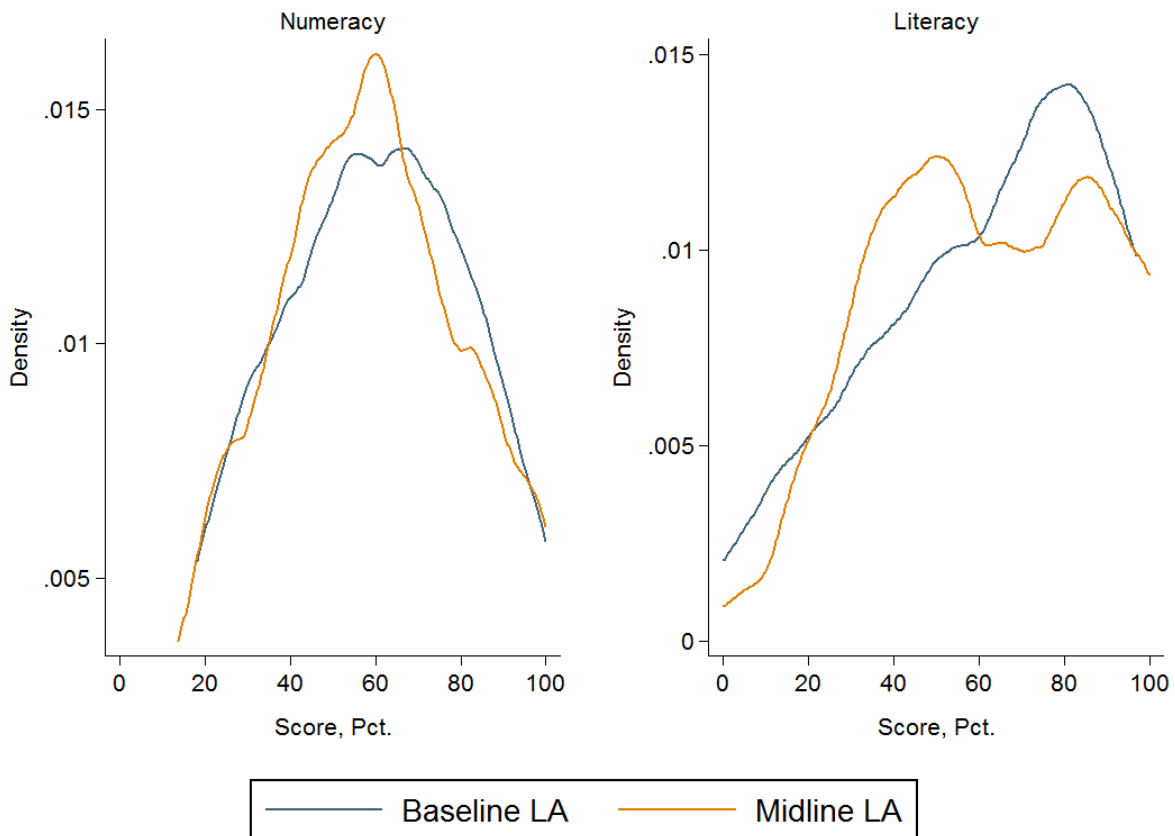
To ensure that our conclusions regarding equivalence were valid, we performed a number of robustness checks. First, we performed alternative significance tests, because t-tests are most adept at highlighting a "location shift" in the distribution of a variable, rather than changes in the shape of the distribution, and may be particularly prone to influence from outliers. We plotted the distribution of baseline and midline scores and examined their distributions for notable differences, as shown in the figure below. We performed a Kolmogorov-Smirnov test for differences in the distributions, and obtained null results (i.e. no significant difference). We also performed nonparametric tests that rely on rank statistics (e.g., the Wilcoxon rank-sum test) and found no significant difference between the baseline and midline numeracy or literacy assessments. Based on the results of the calibration exercise, no equating coefficient was applied to adjust for differences in difficulty between the baseline and midline learning assessments.

---

<sup>331</sup> Girls in this subsample are also part of the broader cohort girl sample. The subsample was drawn from among targeted cohort girls.



Figure 41: Distributions of learning scores for calibration girls



### Practical Challenges in Evaluating Learning

A primary challenge regarding data cleaning (which affects learning results) was the presence of girls in the dataset who had reported midline grade-levels that were unexpected (i.e. outside of the range of grades 3-6 (4-7) that were identified for evaluation at midline 1). The presence of grade 3 (3) girls in the dataset was deemed plausible due to the fact that some girls who were in grade 3 (3) at the baseline were asked to repeat grade 3 for reasons that could be successfully verified in most cases. On the other hand, the presence of girls reporting that they were in grade 1 or grade 2 at the midline was neither plausible (on the face of it) nor verifiable. These girls with unverified grade levels have been excluded from all analysis related to measuring changes in learning over time.

Additionally, the complexity of conducting learning assessments – relative to typical surveys – led to a small number of errors in implementation. The most important instance of this, by far, was on the subtask 4 of the literacy assessment, which required enumerators to measure the time it took girls to reach a given passage. One enumerator misunderstood the directions for this subtask early in fieldwork and did not properly record the number of words read for several of their initial learning assessments. In the six affected cases, literacy scores were calculated in two separate ways: in the first method, we dropped these six observations altogether, not including the respondents in our analysis at all; in the second method, we coded their responses to subtask 4 as missing, and calculated their literacy score using the

remaining 7 subtasks only. In practice, the effect on the main results is negligible, and all results that we tested against both versions of the data were robust to the use of either method.

## **Annex 15: Sampling Framework**

Provide updated and final excel file. The final selection of the schools/communities for the evaluation should be clear.

