

Project Evaluation Report

Report title:	Endline Evaluation of the GEC-T Excelling Against the Odds Project
Evaluator:	One South
GEC Project:	Excelling Against the Odds
Country	Ethiopia
GEC window	GEC-Transition
Evaluation point:	Endline
Report date:	May 2021

Notes:

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



External Evaluation Report

For the Girls' education Challenge Excelling Against the Odds Project, implemented by ChildHope UK and CHADET in Ethiopia

By Tariq Omarshah, Andres Navarrete, and Shiran Anthony de Silva from One South, LLC

May 2021



ONE SOUTH

ONE SOUTH, LLC.

1521 Concord Pike #301
Wilmington, DE 19803
United States of America

☎ +1 703 584 4081

WWW.ONE-SOUTH.ORG

Table of Contents

1. Executive Summary.....	4
1.1 Project Overview.....	4
1.2 Evaluation Purpose and Design	5
1.3 Evaluation Findings.....	7
1.3.1 What impact did the project have on target outcomes?	7
1.3.2 What worked and did not work to improve target outcomes? ..	14
1.3.3 How relevant was the project to the needs and conditions of key groups? ..	18
1.3.4 How sustainable are project achievements?	20
1.4 Conclusion	22
2. Impact & Effectiveness	25
2.1 Attendance	25
2.1.1 What impact did the project have on attendance outcomes? ..	25
2.1.2 What worked and did not work to improve attendance outcomes? ..	32
2.2 Life Skills.....	36
2.2.1 What impact did the project have on life skills outcomes?.....	36
2.2.2 What worked and did not work to improve life skills outcomes? ..	40
2.3 Parent/Caregiver Attitudes & Engagement	43
2.3.1 What impact did the project have on parent/caregiver attitudes and engagement?	43
2.3.2 What worked and did not work to improve parent and caregiver attitudes and behaviours?.....	47
2.4 Teaching & Learning	48
2.4.1 What impact did the project have on teaching quality and learning outcomes?	48
2.4.2 What worked and did not work to improve teaching quality and learning outcomes?	57
2.5 Transition	61
2.5.1 What impact did the project have on transition?	61
2.5.2 What worked and did not work to improve transition?	66
3. Relevance	70

3.1	How well did the project reach marginalized girls?.....	70
3.2	How relevant was the project to the needs and conditions of target groups?	71
3.3	To what extent did the project promote gender equality and social inclusion (GESI)?	79
4.	Value for Money	81
4.1	At what cost did the project deliver impact on key outcomes?	81
5.	Sustainability	83
5.1	How sustainable are project achievements?	83
6.	Recommendations	86
	The following recommendations are made to inform project design and implementation in project regions in the future:.....	86
	Annexes	90

1. Executive Summary

1.1 Project Overview

The Excelling Against the Odds project supported 16,481 marginalized girls, across 27 secondary schools and 47 primary schools in Ethiopia. The project was implemented in partnership by ChildHope UK and CHADET in two zones of the Amhara Region, South Wollo and South Gonder, and one zone of Oromia Region, Arsi, from 2017-2021.

The project aimed to enable marginalized girls to learn in school and successfully transition within school, to vocational training, university or to employment. To achieve this, the project addressed key barriers to girls' education in Ethiopia through teacher professional development, the provision of supplementary learning, improving safety, parental outreach, and other school and community-based activities.

The Girls' Education Challenge (GEC) was launched by the UK Government in 2012 as a commitment to provide 12 years of education to the most marginalized girls in the world. The GEC currently supports girls' education projects in 17 countries. This project was funded through £4,703,652 provided by the UKAID GEC Transition (GEC-T) Fund, the second phase of the GEC.

The GEC-T phase of the project ran from 2017-2021 and was preceded by a first phase, which ran from 2013-2017. The first phase of the project was funded through the GEC Step Change Window and represented an investment of £3,800,000.

The first phase of the project supported girls in primary school to learn and access school. The second phase of the project supported the same cohort of marginalized girls as they entered adolescence, secondary school, and continued to vocational training, university and employment.

In March 2020, the project made several adaptations to implementation in response to school closures as a result of the COVID-19 pandemic. This included providing teaching and learning materials, to support girls to continue learning at home, offering psycho-social counseling support to girls who needed it, and disseminating information, soap, and hygiene materials to prevent the spread of COVID-19.

1.2 Evaluation Purpose and Design

This report summarizes the findings of a multi-year external evaluation. The evaluation appraises the project against the agreed criteria of relevance, impact, effectiveness, sustainability, and value for money. The evaluation team is external to the project and has provided an independent assessment and review of evidence against these criteria to draw evaluation conclusions.

This evaluation report responds directly to 4 central evaluation questions that have been agreed in consultation with the project team and the GEC fund manager (FM).

The evaluation draws from data collected across three evaluation periods: a baseline conducted between May 2017 and December 2017¹, a midline conducted between October and November 2019 and the endline, which draw from data collected between December 2020 and February 2021.

The baseline study was conducted by the project's former external evaluator. The midline and endline evaluations were conducted by the current external evaluator (EE), One South, LLC. The midline evaluation was conducted 2 years after the baseline, as intended, while the endline was conducted 13 to 15 months since the midline, due to delays caused by COVID-19 pandemic school closures and subsequent access challenges.

The evaluation used mixed methods, employing both qualitative and quantitative approaches to gather evidence to answer the evaluation questions.

Quantitatively, at endline the study tracked a cohort of 818 marginalized girls and their caregivers in project areas (410 treatment) and non-project areas (408 control). At Endline, each of these tracked cases includes data from midline and endline including data collected through i) Girls Surveys, ii) Head of Household & Caregiver Surveys, iii) historical attendance tools, iv) learning assessments and v) historical grade tools.

At baseline and midline learning outcome data was collected relying on tools developed following the structure of the Early Grade Reading Assessment (EGRA) and the Early Grade Mathematics Assessment (EGMA). For girls in secondary grade levels the study developed secondary level assessments following guidance provided by the GEC fund manager.

No learning data was collected at endline due to the ethical concerns particularly with regards to the excessive burden assessments would place on marginalized girls so soon after schools re-opened following closures caused by the COVID-19 pandemic. Additionally, the project decided there was little

¹ Baseline data was collected by the previous external evaluator over an extended period due to challenges with field work.

utility in measuring children’s learning levels due to the presumed loss in learning caused by the lack of school for over 6 months.

To understand learning impact at endline, internal grade data for the tracked cohort and national exam data at the aggregate school level was collected and analyzed for representative samples of girls in non-project schools (control) and project schools (treatment).

In addition to the main tracked cohort, the study also conducted teacher surveys with teachers trained by the project, principal surveys with principals in project and non-project schools, and a district official survey with woreda officials.

In-depth interviews and focus group discussions were held in project areas as part of the qualitative research (phase 1), and sampled 35 teachers, 10 school leaders, 20 parents, 26 girls, 4 woreda officials at the district level and other project stakeholders. The evaluation team also conducted additional key informant interviews and discussion groups with project team members.

Findings from the draft evaluation report were shared with key project stakeholders as part of a validation workshop held in March 2021. Participants included school leaders, woreda education officials, and project field and community officers.

Accounts from the validation workshop, including areas where stakeholders brought forward supporting, explanatory, or contradictory views on given evaluation findings, have been documented throughout this report, alongside relevant findings.

Throughout the report several statistical analyses are conducted to support the evaluation team in fully answering the research questions. A list of relevant tests has been provided in the footnote to support access for a non-technical reader².

Annex 1 contains more information on the project design. Annex 2 contains additional details on the evaluation approach, sampling, and methodology.

² Chi-square tests are used to test for associations between two-level categorical variables. Typically, this is used in the report to assess associations between group membership and evaluation status (treatment/control). Independent Samples t Tests are used to compare the means of two independent groups in order to determine whether there is statistical evidence that the associated population means are significantly different; Dependent Samples t-tests are used to compare the means of the same group between periods. Linear regressions modeling is used to test whether a given independent variable predicts a given dependent variable (typically an outcome) under given conditions. Logistic regressions are used to test whether a given categorical independent variable predicts the likelihood of being categorized in a given group for a particular dependent variable. These are used where outcomes are categorical rather than continuous. Statistically significant results are reported at the 5% level ($p < 0.05$) as per commonly accepted practice. This means there is less than a 5% chance of incorrectly rejecting the null hypothesis of a given statistical test.

1.3 Evaluation Findings

1.3.1 What impact did the project have on target outcomes?

The project had a statistically significant impact on Grade 8 Primary School Exam results between 2017 and 2019.

The evaluation gathered Grade 8 exam results for a representative sample of girls in each of the last three exam cycles (Ethiopian Calendar 2010, 2011, 2012) from schools in project and non-project areas across all three years. This included a sample of 1,511 girls in non-project areas (control) and 1,514 girls in project areas (treatment)³.

Grade 8 examinations are managed through the National Educational Assessment and Examinations Agency and are standardized examinations covering the following subject areas: local language (Amharic or Afaan Oromo), English, mathematics, physics, civics and social studies, biology and chemistry.

A cross-sectional impact model finds that the interaction variable⁴ was a statistically significant predictor of overall exam results, when looking at the period between the first and last exam cycle gathered (2017 and 2019)⁵.

Mean results indicate that whilst on average exam results in control schools decreased between 2017 and 2019⁶, exam results in project schools stayed largely the same, due to the project's support⁷.

When we examine each year independently, the project did not have a visible impact on overall exam results between 2017 and 2018.

However, between the last two exam cycles (2019 and 2019), a cross-sectional linear regression model finds the interaction variable to be a statistically significant predictor of overall scores and that girls in project areas had an average change of 5.7% in exam score greater than the change experienced in non-project areas.

The evaluation also gathered data for three subject areas of the exam: local language, English, and mathematics. Examining results for each of these

³ 2010 exam data: 375 G8 girls in control and 377 in treatment; 2011 exam data: 378 G8 girls in control and 378 in treatment; 2012 exam data: 380 G8 girls in control and 381 in treatment; aggregate populations compared between periods using cross-sectional approach with time, treatment, and interaction predicting outcomes levels

⁴ The product of time and treatment status as per cross-sectional impact methodology to isolate effects of the intervention

⁵ $p < 0.05$; Beta=6.072; R square = 0.058

⁶ From a mean of 56% in non-project areas to a mean of 50% in non-project areas

⁷ Mean scores for schools in project areas held steady at 56% between the two years

subjects suggest that the project's impact on overall exam results was driven largely by impact on the English literacy section of the exam.

Between 2018 and 2019 for example, whilst girls in grade 8 non-project schools on average decreased their English literacy exam scores by an average of 6%, in project schools there were no statistically significant differences in girls' grade 8 English literacy exam results between 2018 and 2019.

Learning assessment data and internal grade data further suggest that the project has contributed to improvements in English literacy.

Due to limitations in baseline data in English literacy, the evaluation could not determine the project's impact on English literacy levels between baseline and midline (2018 to 2019)⁸. However, there are several indications that the project supported and contributed to improvements in English literacy.

At midline, in 2019, the evaluation identified significant differences in English proficiency levels across a range of subtasks between girls who were supported by the project and girls in non-project areas⁹. For example, tests for association at this period, find that girls supported by the project are less likely to be categorized as non-readers¹⁰ in English literacy than girls in non-project areas, at statistically significant levels.

Furthermore, a review of school grade level data between Semester 1 and Semester 2 of the Academic Year 2018-2019, finds that the project had a statistically significant impact on English literacy, as measured through semester grades for English.

The impact model, following a difference-in-difference design, was significant and indicated that the project contributed to a 2% increase in English semester grades, *over and above changes* that would have been experienced without the project.

⁸ English literacy at baseline was collected for girls in Grade 7 and Grade 8 on a minority of the overall sample. In addition, an in-depth review of literacy data collected by the previous evaluator at Baseline identified several inconsistencies which were corroborated by reports from enumerators who participated in the field work, relating to test administration. Thirdly, at Baseline, both literacy and numeracy results suggested a counter-intuitive trend: namely that as grade level increased both numeracy and literacy levels decreased or did not change. In consultation with the Fund Manager at Midline, and after a review of Midline learning data to compare results, the evaluator, the project, and the fund manager agreed that utilizing incomplete and limitations in reliability in literacy data from baseline would not meaningfully address the evaluation questions. Therefore literacy data at baseline was not integrated into the analysis.

⁹ This was true for all subtasks reviewed: oral reading fluency, basic reading comprehension 1, advanced reading comprehension 1 (written), and advanced reading comprehension 2 (written)

¹⁰ Non-readers read at a fluency rate of less than 5 words per minute were classified as non-readers.

This is particularly notable because impact was detectable in the short window of a single semester¹¹. In addition, the magnitude of the change experienced by project girls who improved their English grades was greater compared to non-project girls. Of the 64.8% of project girls who increased their English literacy levels between semester 1 and 2, 67.9% increased these levels by 5% or more compared to only 46.8% of non-project girls who improved their English literacy grades in this period¹².

The project had a statistically significant impact on numeracy outcomes between 2018 and 2019.

The project had a statistically significant impact on numeracy outcomes between baseline (2018) and midline (2019)¹³, based on the cross-sectional difference-in-difference model.

Girls supported by the project at midline, on average, had a 12.5% higher score on numeracy¹⁴ compared to baseline girls in both non-project and project areas and non-project girls at midline¹⁵.

Furthermore, at Midline, in 2019, across grade levels, and examining a wide range of numeracy sub-skills, a higher proportion of girls in schools supported by the project met expected curriculum competencies than girls in non-project schools.

The greatest change in numeracy outcomes between baseline and midline was experienced by girls supported by the project who were in grade 8 at the start of the GEC-T phase. These girls would have been exposed to the greatest number of years of the project's first phase, likely exhibiting sustained benefits from prolonged exposure to activities and support.

There is little evidence that the project made a significant contribution to local language literacy.

Due to limitations in literacy data at baseline¹⁶ impact on local language literacy between baseline and midline periods could not be assessed.

¹¹ Comparisons between years were not well-suited to describe changes as grades are likely incomparable between teachers. However, comparisons of internal grade data in the same academic year are based on the same teachers' grading practices.

¹² Though this measure of academic achievement (grades) is non-standardized and based on internal assessments, this finding gives an indication that the project has contributed to improved English literacy outcomes, in the absence of other data sources.

¹³ Numeracy at baseline and midline was measured through a written subtask completed by tracked girls in project and non-project schools, assessing a range of numeracy sub-domains; A wider range of numeracy subtasks were also measured at Midline

¹⁴ See previous footnote

¹⁵ Due to the cross-sectional nature of the model as the previous evaluator didn't collect individual tracking information matched to learning data, the specific proportion of improvement attributed to the project between baseline and midline cannot be estimated as it could be in an individual level tracked design

¹⁶ See footnote 9

At Midline, there are statistically significant associations for being a non-reader in local language literacy and being in a non-project area. However, the evaluation cannot determine whether this reflects gains made in GEC1 or GEC-T. In addition, unlike with English literacy where similar associations were observed, there is limited evidence of detectable impact on local language literacy as measured through semester grades in Afaan Oromo or Amharic.

In interviews with teachers and girls', some participants discuss how the project had supported them access to local language reading materials. Quantitative evidence suggests that some specific activities supported existing local language literacy levels but that no activities were found to support local language learning improvements, leading the evaluation to conclude that there is little evidence of the project's contribution to this outcome.

The lack of significant project contribution to local language literacy, may partially be explained by the fact that secondary school girls' have fewer opportunities in the school environment to practice and reinforce local language literacy learning. As English is the language of instruction in secondary school girls practice English literacy within most lessons. In a similar way, mathematics is reinforced in a range of natural science subject areas at the secondary level.

The project contributed to improvements in teaching quality, particularly in promoting improved lesson preparation, assessment, and pedagogy.

Lesson observations conducted in 2019, indicate that the project has supported teachers to adopt improved preparation, assessment, and pedagogic practice in their lessons:

- 46.8% of lessons in project school demonstrated improved preparation¹⁷, compared to 28.8% of lessons in non-project schools.
- 16.0% of lessons in the project schools demonstrated improved pedagogy¹⁸, compared to 5.5% of lessons in non-project schools.
- 38.3% of lessons in project schools had adopted improved assessment¹⁹ practices compared to 34.2% of lessons in non-project schools.

Lesson observations were not conducted at endline after ethical concerns were raised by the project team in consideration of the increased burden that has been placed on teachers since the COVID-19 pandemic and the additional

¹⁷ See teaching quality section of main report

¹⁸ See teaching quality section of main report

¹⁹ See teaching quality section of main report

pressures observations may impose. The evaluation is therefore limited in the extent to which it can comment on changes in practices after midline.

However, teachers trained by the project interviewed at endline explained how improvements in assessment, preparation, and pedagogy work to support teaching and learning. Teacher survey results also suggest significant differences between project and non-project teachers' attitudes towards the adoption of improved practices. This indicates the project has had a role in changing teachers' mindsets alongside observable differences in teaching practices between project and non-project teachers in 2019.

Teachers interviewed at endline reported that the project increased their understanding, in particular, in active learning methodologies, differentiation, inclusion, the importance of continuous assessment, and strategies to promote group and paired work.

The project supported girls to maintain achievements in attendance carried over from the first phase, GEC1.

The project made significant gains in supporting the attendance of marginalized girls in the first phase²⁰, GEC1, and maintained these gains in GEC-T.

However, the difference in difference models to test for project impact between evaluation periods were unable to isolate project impact at statistically significant levels.

In 2017, close to the start of the GEC-T phase, attendance records across a range of months indicate that girls in project schools had relatively high attendance (mean: 91%). The project was able to maintain these high attendance levels through 2018 (mean: 91%) and 2019 (mean: 92%). In contrast, the cohort of girls in non-project schools started 2017 with lower attendance (mean: 53%) and by 2019 (mean: 70%) had made significant gains but had not reached the high levels of attendance seen in project schools.

The project supported parents to improve their attitudes towards girls' education but due to pressures imposed by the COVID-19 pandemic, the amount of time girls spent on chores increased in both project and non-project areas.

By Endline, in 2020, 45.4% of girls' caregivers in project areas had positive attitudes towards girls' education compared to 23.9% in 2019. Attitudes towards girls' education was measured through a multi-item attitudinal scale focusing on several gender barriers to girls' education and parents' perceptions around whether and when to prioritize girls' education over other pressures.

²⁰ Project staff interviews and consultations based on internal review of attendance data; Endline evaluation report for GEC-1 (2018)

There is a statistically significant association between having positive attitudes towards girls' education and evaluation status. according to a Chi-square test. Parents and caregivers of girls supported by the project are more likely to have positive attitudes towards girls' education than parents and caregivers of girls in non-project areas, at statistically significant levels. This finding indicates that the project has contributed towards improved attitudes amongst parents and caregivers in the last year, despite school closures and despite additional pressures imposed by the COVID -19 pandemic.

Interviews conducted with school, community, and project stakeholders indicate that COVID-19 has placed additional burden on girls and households, often resulting in increases in time spent doing chores. According to stakeholders who participated in the validation workshop, as households face additional economic pressure due to COVID-19, girls are expected to perform more household chores such as cleaning or fetching water, perform childcare responsibilities, or contribute to the household income by working for cash or kind.

Between midline (2019) and endline (2020), the time girls spent on chores increased in both project and non-project areas.

In 2019, girls in project areas spent an average of 3 hours and 20 minutes on chores a day²¹. In 2020, girls in project areas spent an average of 3 hours and 26 minutes per day doing chores. This represents an average increase of 6 minutes between years. Girls in non-project areas by comparison experienced an average increase of 15 minutes in chores per day between 2019 and 2020 (from an average of 3 hours 30 minutes to an average of 3 hours 45 minutes).

Whilst at Midline mean differences were not statistically significant, mean differences at Endline are statistically significant indicating that girls supported by the project have a lower chore burden than girls in non-project areas.

The project contributed to improvements in girls' sense of school belonging, their self-esteem, and their academic self-efficacy.

Conclusions on changes in life skills between baseline and midline are limited due to limitations in data available on life skills at baseline²². However, there are several indications that the project supported life skills outcomes and contributed to improvements between these periods.

In 2019, at midline, for all life skills²³, mean differences between girls supported by the project and non-project girls are different at statistically significant levels.

²¹ Measured through questions on the Girls' Survey asked to girls directly about how much time they spend doing chores

²² Life skills that were assessed at baseline were not measured in the same way to subsequent measurement approaches used at Midline and Endline as the scales featured in the Girls Survey were not included at baseline. In the case of self-esteem, multiple items were used to report self-esteem findings at baseline but do not measure the construct of self-esteem.

²³ Self-esteem, academic self-efficacy, and school belonging

In all cases the treatment group had higher mean life skills at midline than the control group. This suggests improvements in life skills may have preceded midline.

Between midline and endline (2019 and 2020), the evaluation was unable to isolate project impact on girls' self-esteem, resilience, academic self-efficacy, or on girls' sense of school belonging. Qualitative evidence from interviews with girls, parents and caregivers suggests that this may be due to role of COVID-19 conditions on suppressing these outcomes.

Although impact on life skills could not be isolated between 2019 and 2020, a review of life skills within the treatment group finds that, girls who are supported by the project exhibited the greatest improvements in their sense of school belonging²⁴. At Midline, 27.3% of girls had a high sense of school belonging compared to 46.6% at Endline.

Girls in qualitative sessions at midline and endline reported improvements in the way they see themselves, their capabilities and in what they think about school. In addition, several project activities were found in the study to support life skills at both midline (2019) and endline (2020).

Collectively, these findings indicate that:

- (1) the project has supported girls to improve their self-esteem, academic self-efficacy, and school belonging.
- (2) most of these improvements likely took place prior to 2019;
- (3) the onset of COVID-19 and school closures likely influenced the projects' ability to improve key life skills between 2019 and 2020 and that changes in life skills between these periods are largely similar between project and non-project girls.

There is little evidence that the project made a significant impact on transition outcomes, although specific activities were found to support transition.

Between baseline and midline and between midline and endline, the project did not have a detectable impact on transition at statistically significant levels. Although there was an increase in transition rates between midline and endline this was largely due to the effects of the national automatic transition policy put in place after re-opening of schools. This policy boosted measurable transition rates in both project and non-project schools confounding the evaluation's ability to isolate project impact on transition if impact had been made.

Internal data gathered by the project after the endline, once girls in grade 8 found out whether they had passed their exams, suggests that a large proportion

²⁴ Changes in school belonging in the control group were largely similar

of girls transitioned between grade 8 and 9. However, this data is only from project schools and improvements cannot be attributed to the project directly, particularly as there is no comparator from non-project areas.

1.3.2 What worked and did not work to improve target outcomes?

Although the project did not have an observable impact on transition, several project activities were shown to support transition outcomes including Girls' Clubs, having received a school uniform, having received a bursary, and having used a reading corner.

Several project activities were found to be associated at statistically significant levels with successful transition. At endline, these activities include receiving a school uniform, being a member of a Homework Tutorial Club, and receiving a bursary from the project. Bursaries were a main intervention activity provided by the project to support transition rates.

At midline, having used a reading corner or having been a member of a girls' club were associated with having successfully transitioned at midline.

Homework tutorials and improvements in teaching quality drove project impact in numeracy and English literacy.

Linear modeling at midline found that participating in homework tutorials predicts higher levels of English oral reading fluency and numeracy at statistically significant levels²⁵. In addition, participating in homework tutorials was a statistically significant predictor of having improved girls' English literacy between semester 1 and semester 2 of Academic Year 2018-19, the period in which the evaluation has been able to isolate project impact on English literacy²⁶. These findings suggest that homework tutorials are an effective means to improve girls learning, particularly in English and numeracy.

Homework tutorials are organized by schools in Ethiopia as part of wider government policy. The project trained teachers who facilitate homework tutorials in specific teaching practices and pedagogical improvements. This approach was chosen by the project to strengthen existing school structures and

²⁵ See midline report for regression analyses; Findings also indicated these activities supported local language literacy levels but as the evaluation was unable to isolate project impact on local language literacy this is not highlighted in the executive summary. However, the projects role in supporting these existing levels is fully detailed in the learning section of the main report (2.4)

²⁶ ($p < 0.05$; Beta=2.065; r square=0.01).

“Homework support tutorials” (homework tutorials) have been held on a weekly basis in project schools since GEC1.

In GEC1, homework tutorial support groups were formed based on pre-assessment results with girls of similar ability levels placed in the same groups.

In the GEC-T phase, homework tutorials continued but initially focused particularly on supporting underperforming girls. However, after feedback from girls, this approach was adjusted to include a mix of girls of different ability levels. This change was made due in order to (1) eliminate stigma facing underperforming girls attending tutorials (2) enable more girls of different ability levels to benefit from tutorials and (3) to better support teachers to develop skills in self and peer assessment, peer support and differentiation.

Girls who participate in homework tutorials and who were interviewed as part of the study appreciated the opportunity the tutorials gave them to ask questions and go over content delivered in the classroom in a different way to normal lessons. Girls also report being supported and encouraged by their homework tutorial tutors which provides a partial explanation of the effects of homework tutorials on key life skills.

Homework Tutorials, school uniforms, having attended a summer school transition camp, and having received financial support from CHADET promoted attendance in schools.

The project was able to maintain high mean attendance levels throughout the GEC-T (higher than 90% for the calendar month examined for all three evaluation periods).

Quantitative evidence indicates that being a member of Homework Tutorials, having attended a summer school transition camp, and having received a school uniform all support attendance outcomes. This was validated through interviews with a wide range of project stakeholders., Stakeholders also emphasized that financial support provided by CHADET through bursaries and the provision of materials played an important role in supporting girls to attend and access school to begin with.

Girls in South Wollo and South Gonder who participated in validation sessions where findings were presented to them also emphasized the role of attending summer school transition camps in supporting them to improve their attendance: *“Before we attended training in summer transition we did not know about life and the challenges we would face in secondary school. But the training supported us to build our capacity and self- confidence on how to cope with these challenges. It gave us an awareness of how to select our friends on how to set a goal and how we will achieve it”*²⁷.

²⁷ Validation workshop girls South Gonder

While attendance levels were already high at baseline in project areas, the study found that attending school more often leads to higher grades in English and mathematics, as measured through internal school semester grade data, validating the project's underlying assumption in targeting this outcome.

Several project activities were effective at supporting girls' school belonging, academic self-efficacy, and self-esteem.

Participating in Girls Clubs²⁸, participating in Homework Tutorials²⁹, having made use of a Reading Corner³⁰, and having received a school uniform³¹ from CHADET each support girls' sense of school belonging at endline.

School belonging is understood as the sense of relatedness or connectedness to one's school and was measured through the School Connectedness Scale³². Connectedness together with autonomy and competence is understood within self-determination theory as key ingredients necessary to engage in goal-directed behaviour³³.

Across intervention activities, participating in a Girls Club and receiving a school uniform were the most effective interventions at supporting school belonging. Exposure to these activities explained the greatest variance in the data compared to other activities.

Participating in Homework Tutorials³⁴, and having made use of a Reading Corner³⁵, supports girls' academic self-efficacy at Endline. Academic self-efficacy is understood as how capable girls feel to participate in and complete academic tasks. Academic self-efficacy was measured through the Perceived Competence for Learning Scale³⁶.

Participating in Girls Clubs³⁷, having made use of a Reading Corner³⁸, and having received a school uniform³⁹ from CHADET each support girls' self-esteem levels at Endline.

Self-esteem refers to the extent to which we like, accept or approve of ourselves and is understood to be a self-evaluation that is also based on how we think other see us. The relationships and experiences that pupils have at school have

²⁸ (p<0.05; Beta=0.116, r square=0.006)

²⁹ (p<0.05; Beta= 0.121, r square =0.009)

³⁰ (p<0.05; Beta=0.097; r square=0.006)

³¹ (p<0.05; Beta=0.143; r square=0.012)

³² Furlong 2011

³³ Self-determination theory is a macro theory of human motivation and personality that concerns people's inherent growth tendencies and innate psychological needs. It is concerned with the motivation behind choices people make without external influence and interference. Research in education utilizing self-determination theory has examined the role of autonomy, competence and relatedness (belonging) in the classroom.

³⁴ (p<0.05; Beta= 0.092, r square =0.007)

³⁵ (p<0.05; Beta=0.068; r square=0.004)

³⁶ Williams, Freedman, & Deci, 1998

³⁷ (p<0.05; Beta=0.084, r square=0.007)

³⁸ (p<0.05; Beta=0.083; r square=0.008)

³⁹ (p<0.05; Beta=0.081; r square=0.007)

been found to influence their development, psychological well-being, self-esteem and social adjustment. Better pupil–peer relationships and also pupil–teacher relationships have been shown to significantly contribute to the development of more positive and less negative self-esteem⁴⁰.

Family hubs and outreach activities conducted by community workers and volunteers to engage parents and caregivers were successful in improving parent attitudes towards girls' education.

Having a member of a household directly approached by CHADET in the last 3 years and spoken to about attendance and enrolment in school was a statistically significant predictor of improved parental attitudes between Midline and Endline. This indicates that the outreach activities are successful in supporting parents to improve their attitudes towards girls' education.

Findings also suggest that being approached by CHADET in the last 3 years to discuss girls' education may have had an influence on the number of hours girls spent doing chores at endline. Differences in mean time spent on chores are different at statistically significant levels between girls whose households had been approached by a community worker and those whose households had not. During outreach visits community workers speak to caregivers about attendance, enrolment, and reducing girls' chore burden.

Supporting girls to feel comfortable participating in class, supporting their self-esteem, and supporting their self-efficacy all play a role in supporting girls learning outcomes.

Supporting girls to feel capable and comfortable participating in class is a statistically significant predictor of English oral reading fluency and numeracy at Midline ($p < 0.05$)⁴¹.

Several activities were found to promote girls' perceived capacity to participate in class including participation in Girls' Clubs and Homework Tutorials. This also suggests that the training teachers have received to increase girl engagement and involvement in class has contributed to improvements in learning.

Several home environments factors relating to caregiver engagement were shown to support learning.

Parental attitudes towards girls' education is a statistically significant predictor of local language oral reading fluency and local language aggregate score at

⁴⁰ Sarkova, M., Bacikova-Sleskova, M., Madarasova Geckova, A., Katreniakova, Z., van den Heuvel, W., & van Dijk, J. P. (2014). Adolescents' psychological well-being and self-esteem in the context of relationships at school. *Educational Research*, 56(4), 367-378.

⁴¹ See midline report for regression analyses

Midline. This is likely due to the fact that most parents don't speak English and are more able to support local language literacy learning, particularly if they have positive attitudes towards girls' education.

Additionally, having an adult at home to help a girl with homework is a statistically significant predictor of English aggregate score and numeracy at midline, and having an adult at home to ask a child about what they do in school is a statistically significant predictor of local language aggregate score, and local language oral reading fluency.

1.3.3 How relevant was the project to the needs and conditions of key groups?

The project made timely and relevant adaptations to support girls to continue learning during school closures as a result of COVID-19.

To support girls in continuing their learning, the project created literacy and numeracy worksheets and homework packs with teachers. These packs contained literacy and numeracy activities for girls to complete. 91% of girls in project areas group received worksheets from CHADET when school was closed due to COVID-19.

To support girls' psycho-social wellbeing the project also offered PSS to girls in need of well-being support which was delivered through club facilitators and tutors. 35.9% of girls in the treatment group have received counseling from an adult over the last few months since returning to school. 100% of those who had an adult counsel them found it helpful.

The project response to COVID-19 effectively equipped girls with improved knowledge of how to prevent transmission and protect themselves and others.

88% of girls in the treatment group report having received a leaflet about COVID-19 from CHADET. 98.7% of those receiving a leaflet found it helped them increase their knowledge of COVID-19 and COVID-19 prevention strategies.

Being in the treatment group is a statistically significant predictor of COVID-19 prevention knowledge levels, suggesting the project has supported girls to know how to prevent COVID-19 transmission.

The project appropriately targeted sexual and reproductive health barriers relevant to girls as they enter adolescent, including early marriage and access to sanitary wear.

Findings from the evaluation validated several assumptions about the sexual and reproductive health barriers to girls' education in the project's context. At midline finding it hard to access sanitary wear, for example, had a negative effect at statistically significant levels on several learning outcomes. The project also appropriately targeted early marriage which, according to school stakeholders and community leaders, has become more prevalent in communities due to the economic consequences of COVID-19 and school closures.

The project has achieved significant results in addressing some of these barriers. Throughout implementation, the project aimed to provide girls with sanitary wear. As part of the first phase of the GEC, the project established sanitary corners that are used as rest and changing areas in schools.

The midline study identified that 9.7% of girls in project schools had difficulty accessing sanitary wear. In response to this, the project worked to improve access to sanitary wear in project schools for all target girls. Over the last year 83.1% of girls supported by the project have found it easier to access sanitary wear, compared to 53.1% of girls in non-project areas. A wide range of stakeholders interviewed report that this has also supported girls to better access and attend school during menstruation. In similar rural settings in Uganda, research further validates this and indicates that improved access to sanitary wear supports attendance outcomes for adolescent girls⁴².

The project appropriately targeted girls living in households facing economic hardship, as several of these factors were found to negatively influence outcomes.

Several findings from the evaluation suggest that specific hardship characteristics such as unemployment, or reduced poverty levels had resulted in reduced learning outcomes. A wide range of stakeholders also report that girls from families facing additional economic burden are more likely to drop out of school or be pressured into early marriage. Significant evidence from the evaluation suggests that interventions providing financial or material support to girls were effective at supporting outcomes validating this assumption.

The project appropriately targeted barriers associated with disability as findings suggest girls with disabilities face additional barriers to accessing, learning, and transitioning in school.

Several findings from the evaluation indicate that girls with functional impairments experience reduced learning and transition outcomes. Interviews

⁴² Keeping African Girls in School with Better Sanitary Care”, ESCR-DFID IMPACT (March 2018)

held with girls with disabilities at different evaluation periods also support this finding. Teachers interviewed also cite that training on how to teach children with disabilities has supported them in areas which they may have been trained in before but have not yet put into practice.

The project appropriately targeted school safety.

4.1% of girls in the project group at endline report that they do not feel safe in school which is largely a similar proportion as at Midline (3.6%). However the evaluation did observe a significant reduction in the proportion of girls who report that they are being bullied between baseline and endline. 4.2% of girls in the project group report being bullied at Endline compared to 9.25% at Midline, suggesting the prevalence of bullying has decreased. Girls in the control group are more likely to be bullied at statistically significant levels according to tests for association, indicating that there are visible differences between project and non-project areas in the prevalence of bullying at endline.

1.3.4 How sustainable are project achievements?

To assess project sustainability, the evaluation organized participatory sessions focused on sustainability with external project stakeholders during the validation workshop. Participants included teachers, headteachers, project officers, community workers, and woreda officials. The session was held after the main validation workshop, so stakeholders were aware of the main findings of all outcomes as they approached these discussions.

Through a participatory process, stakeholders rated the sustainability of project achievements across selected outcomes. Findings on sustainability are based on these consultations and the conclusions agreed by participants.

Attendance achievements are likely to be sustained by well-established supports at the community, school, and system level.

Stakeholders widely agreed that in schools there are several actors and support mechanisms in place to promote attendance outcomes and closely monitor children's attendance. Similar mechanisms were reported at the system level. At the community level however, stakeholders rated support as "emerging". Improvements in attitudes towards girls' education and some behaviors have been identified, barriers to attendance outcomes still exist in communities, particularly those caused by abductions and early marriage. The project disagreed with this rating at the community level. Community workers at the project level will likely also continue to support attitudinal improvements after project closure and this was raised by stakeholders during the project closure workshop.

Learning achievements are likely to be sustained by support that is becoming established at the system level, and emerging supports at the school level.

Stakeholders indicated that there was strong commitment and support to improve learning outcomes at the system level and even expected budgetary commitments to do so in line with wider government policies.

At the school level, support was rated as emerging as although there have been improvements in learning outcomes and schools are interested for these to continue, stakeholders argued that the consistency and quality of the homework tutorials are dependent on financial support provided by CHADET. While all schools provide tutorials as per government policy, the small transfers provided by the project supported teacher motivation and consistency in delivery, according to stakeholders.

At the community level, stakeholders could not identify specific community support that directly target girls' learning outcomes. While this is outside of the theory of change, it is important to note with regards to the sustainability of achievement made through this project. In conditions in which schools are closed, for example, community supports to sustain learning such as improved access to books, or access to appropriate community-based supplementary learning, could play a role in sustaining outcomes.

Transition achievements are likely to be sustained by emerging support at all levels.

Stakeholders reported that parents and caregivers have changed their attitudes to support girls to enroll and continue in school, although barriers still exist, particularly economic barriers that limit their ability to continue in school. At the system level, participants report that although there is a commitment to support girls to transition there is a need to identify best practices that can be used for replication which would require gathering more data.

Teaching quality improvements are likely to be sustained through support that are becoming established at school level and emerging support at system level.

At system level, although support for continued teacher professional development is a part of on-going sector activities, there was some concern amongst participants that this would not be at the same level of effort as provided by CHADET. At the community level, participants report that there is little awareness from caregivers as to the importance of teaching quality, signaling that there may be a lack of demand for teaching quality improvements in the future from parents and caregivers.

Life skills improvements are likely to be sustained by emerging support at the community level and support becoming established in schools.

Stakeholders widely agreed that schools were well aware of the importance of each of the life skills reviewed including self-esteem, school belonging, and academic self-efficacy and are actively considering initiatives to support these outcomes in the future.. At the community level, stakeholders could not identify specific community support that directly target girls' life skills outcomes. Project staff have realized that the CCCs at community levels will likely support some life skills achievements after project closure. At community level stakeholders agreed there is an increasing awareness of the importance of these life skills for girls, particularly through outreach activities in family hubs. At the system level, however, stakeholders report that there is little awareness on the role of life skills or their relative importance to educational outcomes. Further engagement with system level stakeholders will therefore need to take place on uptake of these findings.

1.4 Conclusion

The Excelling Against the Odds project addressed relevant barriers to girls' education in target areas and was able to make timely, relevant, and effective adaptations in the face of changing conditions due to the COVID-19 pandemic. These adaptations were able to support girls to continue learning despite school closure. In addition, project messaging on COVID-19, resulted in increased knowledge amongst girls and communities of prevention and mitigation strategies, including the use of social distancing, mask wearing and handwashing.

Several intervention assumptions about what was expected to work to improve target outcomes and about the barriers that influence target outcomes were validated by the external evaluation. The evaluation confirmed that economic hardship, school safety concerns, early marriage, poor sexual and reproductive health, and experiencing a disability have a negative influence on quality education outcomes in project regions.

The project successfully delivered improvements in key outcomes including impact in English literacy and numeracy, improvements in teachers' mindsets and practices, and improvements in parental attitudes towards girls' education. Several activities were also shown to support girls' school belonging, self-esteem, and academic self-efficacy.

Impact on learning outcomes was driven particularly by the success of homework support tutorials delivered by tutors trained by the project and through improvements made in teaching practices, including in lesson

preparation, assessment, and pedagogy. Activities focused on homework support tutorials strengthened existing school delivery of supplementary learning for children and was also shown to support girls' comfort levels participating in class.

Evidence on what works indicates that:

- (1) Supporting girls to feel comfortable participating in class supports English literacy and numeracy outcomes
- (2) Homework tutorials delivery and improvements in teachers' preparation, assessment, and pedagogy are effective mechanisms to improve numeracy and English literacy outcomes
- (3) Home environment factors including parental engagement and positive attitudes support learning outcomes
- (4) Bursaries, school uniforms, and girls' clubs support transition outcomes
- (5) Self-esteem supports English literacy and numeracy outcomes and Girls Clubs, reading corners, receiving a school uniform all support girls' self-esteem
- (6) Participating in a Girls Club and receiving a school uniform were the most effective interventions at supporting school belonging.
- (7) Academic self-efficacy supports English literacy outcomes. Homework tutorials and reading corners support girls' academic self-efficacy.

Based on data available, there is little evidence the project had an impact on local language literacy levels or on girls' transition rates.

Findings suggest that transition rates are high in both project and non-project areas, although the automatic transition policy carried out in Ethiopia after school re-openings may have prevented the meaningful measurement of project impact on this outcome. Internal grade 8 project data suggests that a high proportion of girls in project schools passed their exams and will progress to secondary schools. However, without comparable data from non-project areas it is difficult to draw conclusions on the extent to which this can be attributed to the project. Furthermore, while the project initially expected some girls to continue to vocational training and employment, no girls in the tracked cohort pursued these transition pathways.

Local language literacy is less relevant to girls as they progress in school, as the language of instruction is English. However, local language literacy is a core skill for girls to access and fully participate in the Ethiopian workforce or continued education and is targeted by the national curriculum.

The project should consider whether either of these outcomes (transition and local language literacy) are relevant to target in future programming and to

further investigate conditions affecting these outcomes, particularly for girls in upper secondary school.

The participatory sessions on sustainability with project stakeholders, including district officials, held in the three regions, indicated that the strongest support for the sustainability of project achievements across outcomes is present at the school level. Schools in general have existing mechanisms or are putting in place mechanisms to sustain project outcome achievements.

Across outcomes, sustainability at the system level is emerging with additional support necessary to increase awareness of the importance of life skills to system level stakeholders.

At the community level, for some outcomes including attendance and transition supports are emerging. However, there is little evidence of supports in place at the community level to sustain learning outcome achievements, such as specific resources available to girls in communities, or community-based learning activities in which they could participate. Additionally, there is little demand from parents and caregivers, despite greater awareness, for initiatives that would support girls learning and continued improvements in teaching quality.

Recommendations based on evaluation findings are listed in Chapter 6.

2. Impact & Effectiveness

2.1 Attendance

2.1.1 What impact did the project have on attendance outcomes?

The project aimed to reduce barriers preventing marginalized girls from attending school. During the GEC-T phase of implementation, the project:

- Set up 74 sanitary corners in schools and provided sanitary pads to girls
- Established 121 Letter Link Boxes (LLBs) in schools to support child safeguarding⁴³
- Trained 43 Community Care Coalitions (CCCs) and school boards to support child safeguarding in schools, families, and communities
- Trained students (boys and girls), school counsellors, principals, focal teachers, and education bureau officials to support girls sexual and reproductive health
- Tracked and traced girls through community workers and community volunteers and engagement of schools to identify girls who have dropped out or who are at-risk of drop-out for specific outreach targeting
- Provided assistive devices to girls with disabilities
- Paid a proportion of secondary school costs and accommodation for the most marginalised girls.
- Supported girls with transport to Sexual and Reproductive Health Services
- Purchased school uniforms and school materials for marginalized girls
- Supported girls' clubs where advocacy and communication activities support girls in managing negative parent, caregiver, and community member attitudes and behaviours towards girls' education
- Provided life skills training for girls and boys

⁴³ Girls can report incidents in a box that is centrally located in the school grounds, which are gathered on a weekly basis by trained focal teachers for appropriate action and response.

- Supported girls with fees to access TVET and transport to and from TVET

The project made significant gains in supporting the attendance of marginalized girls in the first phase, GEC1, and these gains were maintained in GECT⁴⁴.

Attendance data at Endline was collected from historical school records for the periods September 2017, 2018 and 2019 using the attendance tool. Attendance from Midline was collected from historical school records for December in South Wollo and South Gonder and November in Arsi for 2019 and 2019. Months were chosen based on consultations with the project team to ensure the academic month selected did not have term breaks and represented a girls' typical attendance levels, accounting for seasonal fluctuations due to weather and harvesting, for example. Comparisons between periods were made between the same calendar months to ensure comparability between measures.

The GEC1 evaluation found that the project had an impact on attendance outcome between the GEC1 baseline and midline but not between the GEC1 midline and endline ⁴⁵. By the endline of GEC1, the control group had mean attendance rates of close to 90%. Unlike the treatment group, these levels were not maintained between 2017 and 2018, given that average attendance levels decreased to a mean of 53% in control by the time of the GEC-T baseline⁴⁶.

During the current GEC-T phase, at baseline, on average, girls in treatment schools attended school 91% of the days that the school was open, in a typical calendar month⁴⁷ compared to an average of 52% in control schools⁴⁸. Midline findings looking at different calendar months also identified a similar trend.

Across all three evaluation points of GEC-T, average attendance rates for girls in the treatment group were significantly higher than those in the control group.

The difference in means between both evaluation groups were statistically significant at baseline⁴⁹, midline⁵⁰, and endline⁵¹.

⁴⁴ No visible additional impact on attendance could be made in the second phase of the project (GEC-T) due to maintained high levels of attendance throughout

⁴⁵ Fujiwara et al. (2017). FROM EARLY MARRIAGE, RISKY MIGRATION, DOMESTIC WORK AND STREET LIFE INTO TRANSFORMATIVE EDUCATION. p.47.

⁴⁶ The GEC1 evaluation was conducted by a previous external evaluator and the data has not been verified by the present author

⁴⁷ This was carried out in September 2018

⁴⁸ Attendance records from baseline for both evaluation groups show an already existing gap between the control group ($M=0.53$, $SD=0.47$) and the treatment group ($M=0.91$, $SD=0.22$) before the project intervention

⁴⁹ $t(583.84)=14.68$, $p=9.53E-42$

⁵⁰ $t(641.55)=10.36$, $p=2.26E-23$

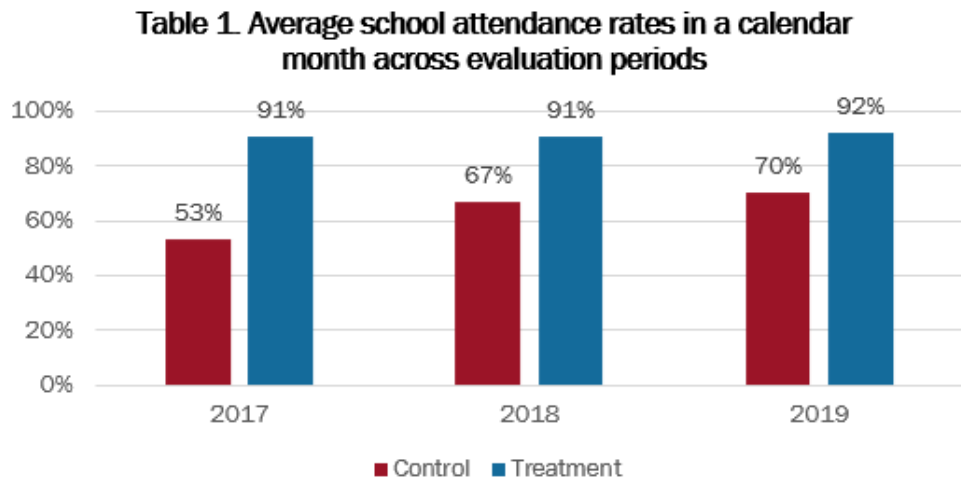
⁵¹ $t(637.69)=9.49$, $p=4.73E-20$

The DiD models that test if treatment status predicted attendance improvements between evaluation periods in the GECT phase, were insignificant, indicating that the project did not have a detectable impact on attendance between the evaluation points⁵².

However, given that girls in treatment schools had, and continued to maintain high levels of attendance from the start of the GEC-T only limited additional gains could be made. This is due to the project making significant progress in increasing and maintaining high levels of attendance throughout GEC.

Attendance data from internal project monitoring in Q12 and following a tracked cohort of girls found that girls in the cohort had on average high attendance levels (99.61%)⁵³.

These findings support high levels of attendance observed as part of the evaluation. Findings from the evaluation are displayed in the figure following.



Improvements in attendance averages in the control group between evaluation periods exceeded improvements in attendance averages in the treatment group, as girls in control schools had more scope for improvement.

This is due to treatment group attendance rates being consistently high over the three data points.

Mean attendance rates between baseline and midline in the control group are different at statistically significant levels⁵⁴ with an average increase of 26.41% between periods. In contrast, the average attendance rates in the treatment

⁵² Impact regression models across key outcomes are shown in Annex XX

⁵³ Representing an increase from 99.35% in Q9 GEC-T Quarterly Project Report (QPR) Q11. However this increase is not statistically significant.

⁵⁴ $t(407)=-9.16, p=2.57E-18$

group largely stayed the same between baseline and midline (91.4% at baseline and 91% at midline). This is likely due to the fact that attendance levels were already high and there was little room for improvement beyond gains achieved in the first phase of the project.

Between midline and endline, differences in average attendance rates in the control group are statistically significant⁵⁵, with an attendance rate of 70% at endline and 67% at midline. Differences in attendance rates in the treatment group between endline and midline were also significant⁵⁶, with an attendance rate of 92% at endline and 91% at midline.

While findings suggest that improvements in the control group were greater between evaluation periods, average attendance levels in the treatment group were significantly higher at each period.

The project was able to successfully maintain these high rates of attendance throughout evaluation periods. Improvements in attendance rates in the control group over time indicate that external factors such as government policies may have influenced attendance rates overall.

Average attendance rates at all grade levels are significantly higher in the treatment group than in the control group across evaluation periods.

According to monthly attendance records from September 2019, average attendance rates in the treatment group are significantly higher than in control for girls in grade 6⁵⁷, grade 7⁵⁸, grade 8⁵⁹, grade 9⁶⁰, and grade 10⁶¹.

Mean differences between treatment and control using monthly attendance records from September 2018⁶² and September 2017⁶³, were statistically significant for all grade levels, with the treatment group having higher attendance levels for both periods than girls in the control group.

⁵⁵ $t(407)=-7.01, p=1.01E-11$

⁵⁶ $t(409)=-4.03, p=0.0007$

⁵⁷ $t(144.45)=3.41, p=0.001$

⁵⁸ $t(244.23)=3.98, p=7.14E-08$

⁵⁹ $t(91.45)=3.41, p=0.001$

⁶⁰ $t(148.09)=5.15, p=8.05E-07$

⁶¹ $t(19.4)=8.38, p=7.16E-08$

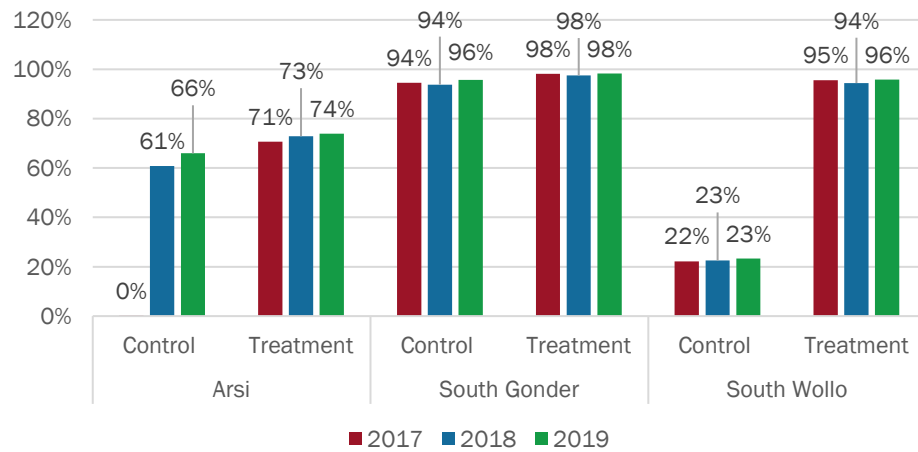
⁶² grade 5, $t(148.58)=3.46, p=0.001$, grade 6, $t(244.5)=4.55, p=7.14E-08$, grade 7, $t(87.62)=4.16, p=0.00008$, grade 8, $t(128.86)=5.01, p=1.08E-06$, and grade 9 $t(35.58)=7.01, p=3.4E-08$.

⁶³ grade 4, $t(130.68)=4.34, p=0.00002$, grade 5, $t(239.21)=5.94, p=1.02E-08$, grade 6, $t(79.84)=8.38, p=1.49E-12$, grade 7, $t(117.83)=7.65, p=5.94E-12$, and grade 8 $t(36.57)=11.83, p=4.61E-14$.

Comparisons between treatment and control groups across grade levels align with overall findings, suggesting that the project may have contributed to treatment group improvements in attendance before GEC-T began.

To understand if there were differences in attendance rates between treatment and control groups across all three project areas, a mean comparison across each evaluation point has been made.

Table 2. Average Attendance Rates by Region by each evaluation point



At baseline, in 2017, mean differences in attendance between control and treatment are statistically significant in South Wollo⁶⁴, in Arsi⁶⁵, and in South Gonder⁶⁶ with higher attendance levels occurring in the treatment group. This supports earlier findings suggesting that improvements in attendance from Phase 1 were carried over into Phase 2.

At midline, in 2018, mean differences in attendance between control and treatment are also statistically significant in South Wollo⁶⁷, in Arsi⁶⁸, and in South Gonder⁶⁹ with higher attendance levels occurring in the treatment group.

At endline, in 2019, mean differences in attendance between control and treatment are statistically significant in South Wollo⁷⁰ and South Gonder⁷¹ only. In Arsi, differences in attendance at statistically significant levels between the control and treatment groups have not been identified which likely reflects a successful wider initiative to improve attendance outcomes in the region, thus supporting these gains in control schools.

⁶⁴ $t(128.72)=18.04, p= 4.71E-37$; treatment ($M=0.95, SD=0.14$), control ($M=0.22, SD=0.4$)
⁶⁵ $t(95.2)=18.62, p= 1.7E-33$; treatment ($M=0.71, SD=0.37$), control ($M=0.002, SD=0.01$)
⁶⁶ $t(374.66)=7.51, p= 4.49E-13$; treatment ($M=0.98, SD=0.04$), control ($M=0.95, SD=0.06$)
⁶⁷ $t(129.03)=17.42, p=1.1E-35$; treatment ($M=0.94, SD=0.14$), control ($M=0.23, SD=0.40$);
⁶⁸ $t(191)=2.31, p=0.02$; treatment ($M=0.73, SD=0.39$), control ($M=0.61, SD=0.34$)
⁶⁹ $t(345.28)=7.38, p=1.22E-12$; treatment ($M=0.98, SD=0.04$), control ($M=0.94, SD=0.06$)
⁷⁰ $t(127.1)=17.18, p=6.03E-35$; treatment ($M=0.96, SD=0.14$), control ($M=0.23, SD=0.41$)
⁷¹ $T(355.85)=5.86, p=9.72E-09$; treatment ($M=0.98, SD=0.03$), control ($M=0.96, SD=0.05$)

With the exception of differences between mean attendance rates in control and treatment in Arsi in 2017, differences between evaluation groups for the region in 2018 and 2019 are less than 10%. In South Gonder, differences in attendance rates are less than 5% across all periods.

In South Wollo, attendance levels between girls supported by the project and girls in non-project areas are statistically significant at all evaluation periods.

In South Wollo, differences between the two evaluation groups remain consistently higher than 70% at all periods, with differences at each period between the two evaluation groups being statistically significant.

These large differences in attendance rates between the two evaluation groups were also identified at baseline. Furthermore, both control and treatment groups see similar increases in attendance rates of 1% between baseline and endline.

The lowest average attendance rates in the treatment group were recorded in Arsi.

Average attendance rates in Arsi in the treatment group (<75%) are lower than those in South Gonder and South Wollo, which consistently remained above 90% across evaluation periods. Differences in attendance rates between Arsi and both South Gondar⁷² and South Wollo⁷³ are statistically significant at each evaluation period.

This suggests girls in Arsi may face additional barriers to attendance outcomes. Recent bouts of civil unrest in the Oromia region may have also contributed towards low attendance rates at endline. Low attendance rates in Arsi were also noted at Midline and will be discussed further in the report's relevance section.

Girls in the validation workshop in Arsi chose to discuss this finding further. They agreed with the finding and explained that this was because parents and caregivers in rural parts of Arsi had not yet changed their attitudes, that many girls wish to migrate to other parts of the country, and this causes them to "neglect" their attendance. They also reported that COVID-19 caused girls in Arsi to lose hope and see marriage as their only life option, and that there is peer pressure to skip school. Additionally, girls reported that those who face economic hardship are pressured to "make money" instead of being in class or to do house chores.

⁷² Arsi & SG 2019: $t(95.64)=-6.0$, $p=3.57E-08$, Arsi & SG 2018: $t(95.9)=-6.23$, $p=1.24E-08$, Arsi & SG 2017: $t(96.2)=-7.25$, $p=1.05E-10$,

⁷³ Arsi & SW 2019: $t(114.6)=-5.14$, $p=0.000001$, Arsi & SW 2018: $t(116.46)=-5.17$, $p=9.78E-7$, Arsi & SW 2017: $t(117.39)=-6.2$, $p=8.71E-9$.

In interviews some girls in Arsi reported finding it difficult to travel to school.

Girls interviewed in Arsi said, *“I travel long distances and there are no friends or students in my village to study together. I am always worried to do my homework alone at home⁷⁴.”*

Another girl in Arsi said, *“the absence of means of transportation makes the journey to and from school very hard... Sometimes we are forced to come back to school to attend tutor class. This increases the frequency of travelling to school in a single day⁷⁵.”*

Despite attendance rates being lower in Arsi, girls in Arsi experienced the largest average increase in attendance between baseline and endline.

With statistically significant differences in means between baseline and endline, $t(95)=-3.92$, $p=0.0001$, Arsi had the largest increase in average attendance rates in the treatment group at 4.2%, in comparison to South Gonder which showed no improvement, and South Wollo which improved by 2.13% (not statistically significant).

Lower attendance rates in Arsi, which remain below 75% across periods, compared to South Wollo and South Gonder, which remain above 90%, suggest that there was more room for attendance rates in the treatment group to improve in Arsi. The control group experienced an increase in average attendance rates in Arsi between midline to endline. Means between baseline and endline were statistically significant, exhibiting an increase of 8.2%⁷⁶.

It is likely that attendance improvements observed in treatment areas are also attributable to other wider government and stakeholder initiatives to improve attendance in the region. This is because the increases observed in the attendance in non-project areas, would signal wider success of other attendance initiatives outside of the projects control. Several government initiatives have focused on supporting school principals to engage with families and caregivers to support attendance of girls.

⁷⁴ AR2: Interview with girl who is enrolled in school and lives far away.

⁷⁵ AR3: Interview with girl who is enrolled in school and lives far away.

⁷⁶ $t(96)=-4.97$, $p=3.02E-06$

2.1.2 What worked and did not work to improve attendance outcomes?

Attending secondary school summer transition camp was a predictor of attendance improvements between baseline and midline.

While controlling for membership in all other project activities, for secondary girls, attending secondary school summer transition camp was a significant predictor of attendance improvements ($p < 0.05$). Attending the camp resulted in girls improving their attendance by 2.3% between baseline and midline.

Secondary school transition camps were organized to support girls after primary school with life skills and transition support training before they transitioned to secondary school. Girls in qualitative session discussed how these trainings supported them to feel more confident and capable attending school, and in some cases reported it had helped them to resist peer pressure to engage in other behaviors.

Differences in attendance rates between girls who participate in project activities and girls who are in control schools are statistically significant across phases. This likely reflects achievements carried over from GEC1 and the success of GEC-T in maintaining these gains.

According to independent sample t-tests, there are significant differences between average attendance levels at endline ($p < 0.05$) of control group and treatment group girls who were exposed to project activities.

Higher averages in attendance were observed in girls who were members or took part in any one of the following activities: (1) Homework Tutorial Clubs, (2) Reading Corners, (3) Girls' Clubs, (4) received school uniform from CHADET, (5) received leaflets from CHADET about COVID-19, (6) received financial support from CHADET, (7) Sanitary Corners for girls 11 and over, and (8) Summer Secondary School transition camp for girls in grade 9 and over.

Although the differences in average attendance levels are significant between girls who took part in each of the activities listed above and girls in the control group, higher attendance rates at endline cannot be attributed solely to these project activities as attendance levels were already at high levels at baseline indicating these improvements took place prior to the GECT phase.

However, high attendance rates in the treatment group ($>90\%$) across all evaluation periods suggests that the work of the project to consistently monitor attendance and track girls at risk of dropping out at community level was a

contributory factor in supporting girls to maintain these initial gains from GEC 1 into and through the duration of GEC-T.

This was also validated by annual household visits conducted by the project at the beginning of the academic year to engage all families, including girls who have not attended school, by the local CCC and the school board.

According to project documents, including quarterly and annual reports, girls in Girls' Clubs, Focal Teachers, and both project Community Workers and Community Volunteers are all involved in identifying girls at risk of drop out and truant girls so that these girls can be immediately counselled and an intervention with their family organized.⁷⁷

As per procedure, the school reports the attendance of girls to Community Workers and Volunteers weekly. This is then triangulated with reports from families, which Community Workers and Volunteers follow up on.

Internal project reporting argues that community Workers and Volunteers, along with principals and teachers, played a strong role by remaining in contact with GEC-T girls during school closures due to COVID-19 which supported the retention rate of 97.89% when school re-opened.⁷⁸

Several stakeholders report that being given a school uniform makes it easier for girls to attend school.

According to one parent interviewed, CHADET has, *“provided them [girls] with education materials like uniform and exercise book. This was very important. It really helped us, especially for a family like me who has been teaching four to five children.”*⁷⁹

Similarly, another parent said, *“Yes, I want to talk about CHADET support. There are many students who have financial problems to buy education materials and uniforms. This is very important.”*⁸⁰

Another girl said that receiving education materials from her tutorials helped her parents change their attitude towards education so that she could keep going to school, *“Even though they [parents] knew how I value education, they were thinking to get me married. Then after I received educational materials, and I convinced them, they now have a good image of my education.”*⁸¹

Teachers also discussed the importance of having a school uniform and being able to attend school, *“The support in uniform is a profound support for both parents and students. [Originally] we had girls who come in months late to*

⁷⁷ GEC-T Quarterly Project Report (QPR) Q8.

⁷⁸ GEC-T Quarterly Project Report (QPR) Q15.

⁷⁹ AR5: Interview with caregiver or parent of girl who is currently enrolled in school and lives far away from school.

⁸⁰ AR12: Interview with girl whose parents/caregivers physically punish them at home.

⁸¹ SG19: In-depth interview with girl who improved her view of herself from HW tutorial.

school because their parents couldn't afford uniforms. The girls also receive exercise books and stationeries⁸²."

Girls report that Girls' Clubs help them attend school more often.

Several girls in sessions reported that the girls club supported them to attend school more often. According to an in-depth interview with a girl on how her opinion on education changed after being a member of a Girls' Club, she said, *"It made me realize what I want to be. And if I am not sick, I don't usually skip school.⁸³"*

During validation workshops with girls in Arsi, several chose this finding and strongly agreed with it. Girls reported that the club gave them a "shared experience with other club members" which made them feel like attending school. Girls also reported that clubs provided them with the opportunity to meet other girls with "better capacity" and to learn from them which made them feel an attachment to school. Finally, several girls in validation in Arsi also reported that the sanitary wear provided to them in clubs helped them to attend school during menstruation.

Girls interviewed as part of the study reported that attending Homework Tutorials helped them improve their attendance.

When asked if attending homework tutorials affect girls' school attendance, one girl said, *"They [girls] don't skip classes often.⁸⁴"*

From a different discussion, another girl said that she had stopped skipping more classes after attending the tutorials, *"before when I was in the previous class, I was skipping classes even if I told my parents I am in the tutorial, but through time I have changed and I became great in classes⁸⁵."*

According to the project, preventing dropouts continue to be a challenge especially in rural and hard to reach communities, but findings suggest that girls from urban areas have lower mean attendance rates.

As noted in the Year 3 Annual Report, project interventions on increasing attendance rates primarily focus on tackling negative social norms, risky migration and the low value placed on girls' education. However, project reporting also suggests that preventing dropouts in rural and hard to reach communities remained a challenge. An increase in the number of dropouts from

⁸² SW12: Focus group with active community of practice that meets regularly.

⁸⁴ SG20: In-depth interview with girl who didn't improve herself regard from HW tutorial.

⁸⁵ SG9: Participatory drawing session on self-regard with girls in HW Tutorial Club

6.24% at the end of Year 2 to 8.69% at the end of Year 3 suggests difficulty in mitigating dropout rates in these areas.⁸⁶

Despite increases in dropout rates recorded by the project between from Year 2 to Year 3, these rates are still below national dropout rates for girls, which are 12% in primary school and 29% in secondary school⁸⁷.

Project reports indicate that as girls become older, dropouts become more likely due to migration to Arab countries or early marriage. Other reasons include a lack of basic necessities, health factors, domestic labor, and parental influence⁸⁸. The project's Management of Drop Out Procedure aims to capture the reasons for dropout and provides steps to be followed in order to respond to these dropouts⁸⁹.

While the project may have highlighted the challenges in reaching rural communities, mean attendance rates across all three evaluation periods were slightly higher for girls from rural areas in comparison to those from either peri-urban or urban areas⁹⁰. However, these differences were not statistically significant. In the control group, statistically significant differences were found between the mean attendance rates of girls from rural and urban areas⁹¹, with higher mean attendance rates recorded for girls in rural areas.

While average attendance rates in the treatment group for girls from either rural or urban areas are high, suggesting that the project has been able to address barriers preventing girls from attending school from either area, findings from the control group suggest that focus may need to be placed on girls from urban areas to ensure that the intervention is fully addressing barriers and that these barriers are well understood, in these areas in particular.

Girls interviewed during qualitative discussions were not aware of the buddy system introduced by the project but think that travelling to school in pairs or groups can positively affect their attendance.

All girls interviewed on issues pertaining to safeguarding at endline were not aware of the buddy system that was introduced by the project to support girls to travel in pairs or groups. The main purpose of this intervention was to increase

⁸⁶ GEC-T Annual Project Report (APR) April 2020.

⁸⁷ Ibid.

⁸⁸ Ibid.

⁸⁹ [CHADET] Management of Drop Out Procedure Dec 2019.

⁹⁰ 2010: Rural Girls $M=0.92$, Urban Girls $M=0.89$

2011: Rural Girls $M=0.92$, Urban Girls $M=0.89$

2012: Rural Girls $M=0.93$, Urban Girls $M=0.90$

⁹¹ 2010: $t(406)=4.93$, $p=0.000001$

2011: $t(323.8)=4.68$, $p=0.000004$

2012: $t(320.31)=4.59$, $p=0.000006$

attendance rates by reducing the risks faced by girls travelling to and from school.

However, while they had not heard of the buddy system, they said that travelling to and from school in pairs or groups would help them feel safe. One girl said, *“You can talk and discuss about school while you travel with friend. It also helps to protect ourselves if something encountered us. The risk is very low⁹².”* Similarly, another girl said, *“Traveling as a pair or in a group would shorten the distance. It also helps girls to attend school on a regular basis. It avoids the risk of encountering abuse or abduction and makes them feel safe⁹³.”*

The onset of COVID-19 may have prevented girls from hearing about the buddy system in target schools. When asked how COVID-19 has affected their safety travelling to school, one girl said, *“...we have been told not to travel or sit together in a group. On the other hand, not travelling in a group affects my safety.⁹⁴”*

2.2 Life Skills

2.2.1 What impact did the project have on life skills outcomes?

The project worked to improve girls’ self-esteem, self-efficacy, belonging and other life skills through several activities. During the GEC-T phase, the project:

- Established 74 Girls’ Clubs where facilitators delivered a life skills curriculum (covering topics relating to safeguarding, leadership, self-confidence, self-esteem, sexual and reproductive health, belonging and other skills)
- Facilitated Good Brothers Clubs with boys to create a healthy and supportive peer environment
- Conducted training on child safeguarding and supported safeguarding initiatives in schools, such as letter link boxes
- Worked with parents to promote support for girls’ education through family hubs
- Trained teachers who facilitate homework tutorials on teaching approaches aimed at supporting girls engagement, academic self-efficacy and self-esteem.

⁹² AR2: Interview with girl who is currently enrolled in school and lives far away from school #1.

⁹³ AR3: Interview with girl who is currently enrolled in school and lives far away from school #2

⁹⁴ Ibid.

- Supported teachers to adopt gender-sensitive teaching practices in schools to create a supportive learning environment

Girls in the treatment group have significantly higher life skills than girls in the control group at Midline, signalling that the project may have played a role supporting life skills between baseline and midline (2018 and 2019) or in the GEC1 phase (2014-2017).

At midline, for all life skills⁹⁵, mean differences between control and treatment are different at statistically significant levels. In all cases the treatment group had higher mean life skills at midline than the control group⁹⁶.

This, in relation to other findings that specific project activities support life skills, suggests that girls may have experienced gains in key life skills between Baseline and Midline. Conclusions on changes in life skills between baseline and midline are limited due to data available on life skills at baseline⁹⁷

For girls clubs in particular, where the project delivers the majority of its life skills content, interviews with club facilitators suggest that most girls were members of the club since GEC1⁹⁸, although the clubs have gained other members in the GEC-T phase.

By endline, differences between treatment and control are no longer statistically significant, indicating that the control group had caught up with the treatment group.

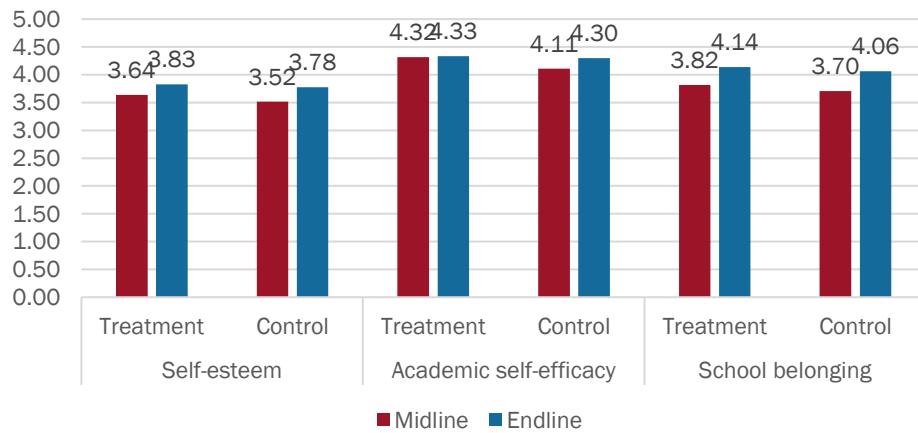
⁹⁵ Self-esteem, academic self-efficacy, and school belonging

⁹⁶

⁹⁷ The previous external evaluation team did not measure life skills at Baseline in a reliable or comparable way to subsequent measurement approaches.

⁹⁸ SG 10 Interview with Club Facilitator

Mean life skills by evaluation group and evaluation period



Between midline and endline (2019 and 2020), the evaluation was unable to isolate project impact on girls’ self-esteem, resilience, academic self-efficacy, or sense of school belonging, possibly due to role of COVID-19 conditions in suppressing these outcomes.

Linear models, using treatment to predict the first difference in each of the life skills outcome were insignificant, indicating that the project did not have a visible statistically significant impact on these outcomes between midline and endline⁹⁹.

Qualitative interviews with school staff, girls, and teachers suggest that school closures and the wider COVID-19 context are factors that have negatively influenced girls’ self-regard, school belonging, and academic self-efficacy and that this may have played a role in suppressing gains in these life skills. This may have suppressed some improvements supported by project activities between midline and endline.

In an interview, one girl described how she felt during the COVID-19 lockdown. She said, *“I did not feel good since staying home ... it was stressful”*¹⁰⁰. She also reported that she missed school as she could only study *“a little bit”*¹⁰¹. Another girl reported, *“It was not a good time, the pandemic made me feel that I am not going to join school again.”*¹⁰²

Parents/caregivers of girls also spoke about how staying at home due to school closures negatively affected their girls. One parent said, *“...they were very disturbed about the lockdown and they did not think that the school would*

⁹⁹ Insignificant in all cases $p > 0.05$

¹⁰⁰ SG17: In-depth interview with girl who improved her view of herself.

¹⁰¹ Ibid.

¹⁰² SG18: In-depth interview with girl who didn’t improve her view of herself.

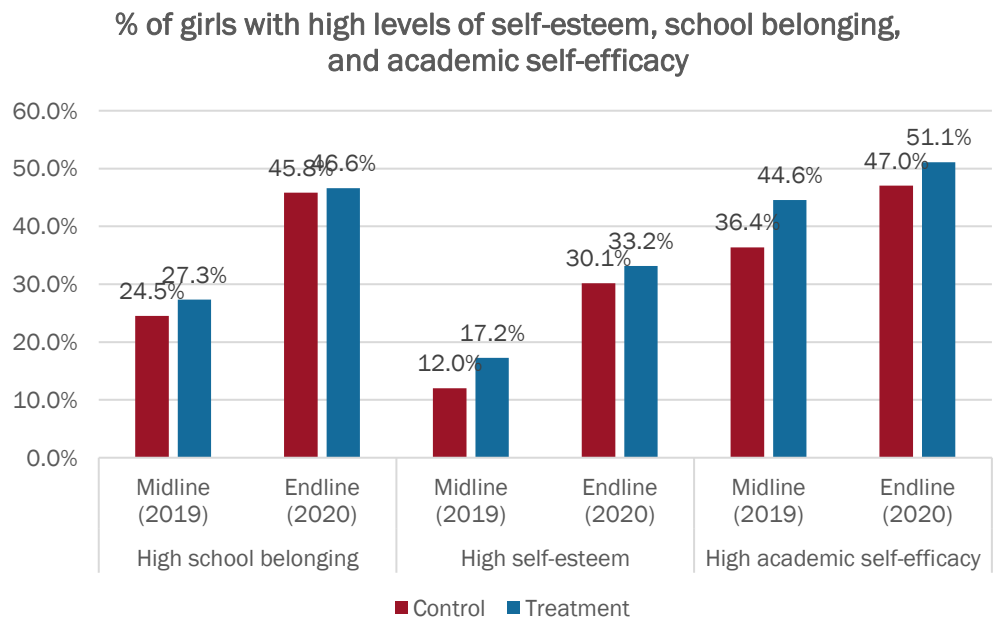
reopen...So they were very scared at that time, and they didn't have time to study due to the workload on the farm."¹⁰³

Between 2019 and 2020, across the Life Skills reviewed, girls who are supported by the project exhibited the greatest improvements in their sense of school belonging.

At Midline 27.3% of girls had a high sense of school belonging compared to 46.6% at Endline. School belonging is understood as the sense of relatedness or connectedness to one's school and was measured through the School Connectedness Scale (Furlong 2011). Connectedness together with autonomy and competence is understood within self-determination theory as key ingredients necessary to engage in goal-directed behaviour, along with autonomy and competence.

It should be noted that similar improvements in life skills were exhibited by girls in project and non-project areas between midline and endline. No differences at endline in life skills levels are statistically significant between evaluation and treatment groups.

No associations were found at endline between treatment status and having 'high self-esteem high academic self-efficacy or high-school belonging.



¹⁰³ SG2: FGD Parents of girls in households facing extreme hardship (at Midline).

2.2.2 What worked and did not work to improve life skills outcomes?

Several project activities are expected to play a role in supporting girls' life skills. These include Girls' Clubs, wider improvements in the teaching and learning environment, Homework Tutorials, and safeguarding activities.

Participating in Homework Tutorials¹⁰⁴, participating in Girls Clubs¹⁰⁵, having made use of a Reading Corner¹⁰⁶, and having received a school uniform¹⁰⁷ from CHADET each support girls' sense of school belonging at endline.

Linear modelling finds that each of these interventions supports school belonging levels at endline at statistically significant levels¹⁰⁸.

By endline the project has provided 61.1% of girls with school uniforms (an estimated 10,070 girls). The project provides girls with school uniforms on an annual basis. Girls are selected by kebele officials based on the agreed criteria. The selection is supported by school principal, community volunteers and community workers.

Stakeholders in the validation workshop corroborated these findings with several reporting the role that girls' clubs play in supporting their life skills. As a participant in validation explained "[Girls Clubs] helped girls to free, develop self-expression, self-confidence and ability to negotiate and succeed in their social as well as academic lives. This helped them to give value for themselves and develop self-esteem".

According to qualitative interviews, Girls' Clubs help girls feel less alone in school.

Several girls reported that Girls' Clubs allowed them to be with their friends. One girl said that she felt lonely at school before joining the club because, "*I had no friends due to poor communication skills¹⁰⁹.*" However, when asked how she felt after joining the club, she said, "*I have friends now so that the feeling of loneliness disappears¹¹⁰.*"

¹⁰⁴ (p<0.05; Beta= 0.121, r square =0.009)

¹⁰⁵ (p<0.05; Beta=0.116, r square=0.006)

¹⁰⁶ (p<0.05; Beta=0.097; r square=0.006)

¹⁰⁷ (p<0.05; Beta=0.143; r square=0.012)

¹⁰⁸ See previous footnotes for regression summary results

¹⁰⁹ SG18: In-depth interview with girl who didn't improve her view of herself.

¹¹⁰ Ibid.

Girls in Homework Tutorials also credited the support they get from CHADET in helping them feel like they are a part of their school.

One girl said, *“Since I’m a part of CHADET, and also a part of regular class in this school, I think I’m definitely a member of the school¹¹¹.”* Other students in the same discussion also voiced similar opinions saying, *“I think I’m a part of it [school] because I got support from CHADET, and I learn in this school¹¹².”*

Participating in Homework Tutorials¹¹³, and having made use of a Reading Corner¹¹⁴, supports girls’ academic self-efficacy at Endline.

Linear modelling finds that each of these interventions supports academic self-efficacy at endline at statistically significant levels¹¹⁵.

Section 2.4.2 (What worked and did not work to improve teaching and learning) examines the relationship between academic self-efficacy and learning outcomes. Homework tutorials more broadly are also discussed in this outcome.

Stakeholders in the validation workshop agreed with this finding. Several explained that during homework tutorials, girls were given awards for high academic performance and that this had a clear role in supporting their perceived academic capabilities.

Qualitative interviews suggest that Homework Tutorials support girls’ academic self-efficacy.

According to one girl when asked how she changed after participating in homework tutorials, she said, *“I was shy before in the regular classes, even if I did not understand, and if I had a question, I would not ask my teachers about it. But after I joined the tutorial classes, I always ask what I did not understand and became a better self-confident student¹¹⁶.”*

Another girl also spoke about her confidence increased after joining homework tutorials. She said, *“Even though I wanted to understand the subject, if the teacher was not explaining well, skipping the concept was the common action. My confidence was low since I did not understand most concepts, but now I am a clever student.”¹¹⁷*

¹¹¹ SG9: Participatory drawing session on self-regard with girls in HW Tutorial Club.

¹¹² Ibid.

¹¹³ ($p < 0.05$; Beta= 0.092, r square =0.007)

¹¹⁴ ($p < 0.05$; Beta=0.068; r square=0.004)

¹¹⁵ See previous footnotes for regression summary results

¹¹⁶ SG9: Participatory drawing session on self-regard with girls in HW Tutorial Club.

¹¹⁷ SG19: In-depth interview with girl who improved her view of herself from HW tutorial.

Participating in Girls Clubs¹¹⁸, having made use of a Reading Corner¹¹⁹, and having received a school uniform¹²⁰ from CHADET each support girls' self-esteem levels at Endline.

Linear modelling finds that each of these interventions supports self-esteem at endline at statistically significant levels¹²¹.

In focus groups girls in South Wollo explained that receiving school uniforms from CHADET made them feel “equal” to their peers¹²². This was also reported by teachers interviewed as part of the study¹²³.

On an annual basis school uniforms are provided to marginalized girls who have been selected by kebele officials based on agreed criteria. The selection process is supported by the school principal, community volunteers and community workers.

Section 2.4.2 (What worked and did not work to improve teaching and learning) examines the relationship between self-esteem and learning outcomes.

53.9% of girls supported by the project are members of Girls' Clubs (an estimated 8,884 girls). Of these girls, 51.6% participate in every club activity, 26.7% in most activities, and 21.7% participate sometimes¹²⁴. 97.3% of girls feel the club facilitator listens to them and acts on what they say.

98.1% of girls find what they learn in the club useful. The most important reason most girls joined the club was to improve their learning in school (59.3% of girls), and the second reason being that they enjoyed club activities (21.3%).

According to qualitative interviews, Girls' Clubs help girls improve their self-confidence.

Several girls reported improvements in self-confidence after participating in the club. One girl explained, “*My self-confidence is respecting people and not being scared of them too. I used to be scared back then and there is nothing scary now. Even if I make a mistake that is how I learn.*”¹²⁵ Another girl in a different discussion said, “*I did not respect myself enough but now I respect myself*”¹²⁶.

Several girls also spoke about how being a member of a Girls' Club gave them more confidence with being more comfortable with their menstruations. One girl said, “*...before I joined the club, I used to think menstruation was a shameful*

¹¹⁸ (p<0.05; Beta=0.084, r square=0.007)

¹¹⁹ (p<0.05; Beta=0.083; r square=0.008)

¹²⁰ (p<0.05; Beta=0.081; r square=0.007)

¹²¹ See previous footnotes for regression summary results

¹²² FGD SW 19 Girls in Grade 8

¹²³ SW 12 COP Focus Group

¹²⁴ 48.9% of girls report the club has too many meetings, 25.3% too little

¹²⁵ SG23: In-depth interview with girl who was in girls club but left girls club.

¹²⁶ SG17: In-depth interview with girl who improved her view of herself.

thing, but now I understand it is part of our nature.¹²⁷” Another girl said, “I feel good about it [joining a club] because they educate me and prepare me for menstruation...speaking about menstruation was not easy. Now I figured it is the nature of all women¹²⁸.

Other stakeholders also spoke about the benefits of Girls’ Clubs on improving the self-confidence of girls. For example, a club facilitator said, “I have seen changes in girls through time like eliminating their fear and develop self-confidence. Before they didn’t have much confidence to come here and ask pads and panties when they have their menstruation in the middle of a class. But nowadays this is changing¹²⁹.”

Furthermore, testimonial evidence gathered from stakeholders to validate these claims also supported these findings. One such stakeholder said, “the life skill training helped them to build their confidence to explain their ideas¹³⁰.”

According to testimonies from relevant stakeholders, receiving a school uniform helps girls with their confidence.

When commenting on what helps girls develop their self-confidence, one respondent in testimonial interviews spoke about how school uniforms help girls. They said, “females get uniform and other material support, so they can be fully confident [of themselves].¹³¹”

2.3 Parent/Caregiver Attitudes & Engagement

2.3.1 What impact did the project have on parent/caregiver attitudes and engagement?

Between 2019 and 2020, the average amount of time girls spent doing chores increased in the control group but remained the same in the treatment group, suggesting the project had a role in preventing similar increases for target girls.

¹²⁷ Ibid.

¹²⁸ SG18: In-depth interview with girl who didn’t improve her view of herself

¹²⁹ SG10: In-depth interview with Girls Club facilitator.

¹³⁰ SG22: Testimonial evidence how girls’ clubs support girls.

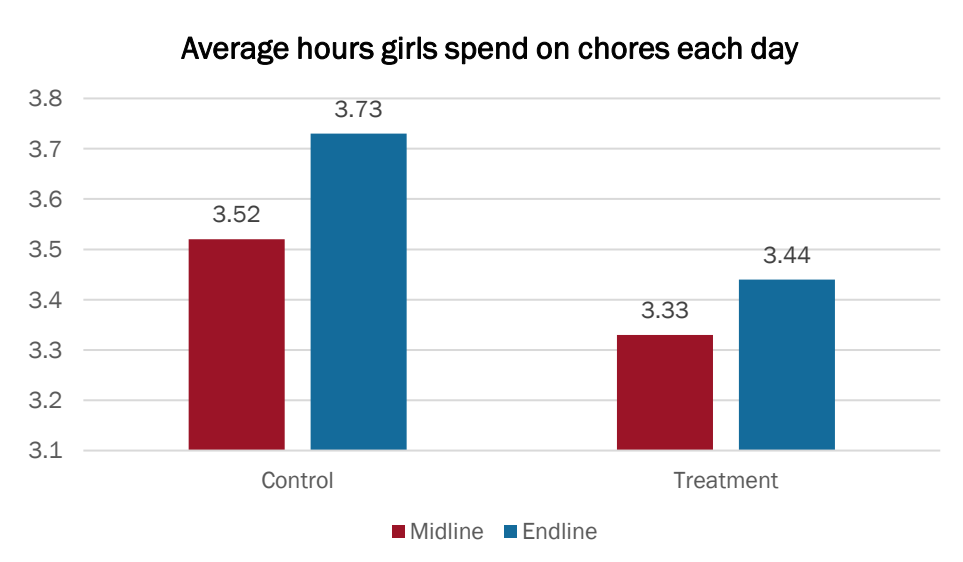
¹³¹ SG24: Testimonial evidence how HW tutorials support girls.

Paired t-tests that compared means of time girls spend doing chores in the control group between midline and endline found statistically significant differences, $t(325)=-2.09, p=0.04$, with a 7.6% increase in the mean number of hours a girl spends on chores.

No significant differences were found between midline and endline in the treatment group. This means that even though both girls in the treatment and control group likely spent more hours at home in the past year due to COVID-19, girls in the treatment group were not subjected to an increase in household chores. This suggests that the project may have contributed towards changing parental attitudes towards the number of chores girls are responsible for and/or increased girls' ability to negotiate the amount of household chores they receive.

On average, girls in the treatment group at endline spent fewer hours on chores each day in comparison to girls in the control group.

Mean hours spent doing chores between treatment and control are different at statistically significant levels at endline¹³², with girls in the control group spending nearly twenty minutes more (3hrs 44 minutes) on chores each day, in comparison to girls in the treatment group (3hrs 26 minutes). However, during the validation workshop some stakeholders reported that the mean of 3 hours per day seemed lower than what they would expect. In some cases stakeholders cited reports of girls doing up to 6 hours of chores a day, especially during school closures due to COVID-19.



¹³² $t(789.23)=-2.54, p=0.01$

Between 2019 and 2020, the project did not have a detectable impact on the number of hours girls spent doing chores or on parent/caregiver attitudes towards girls' education.

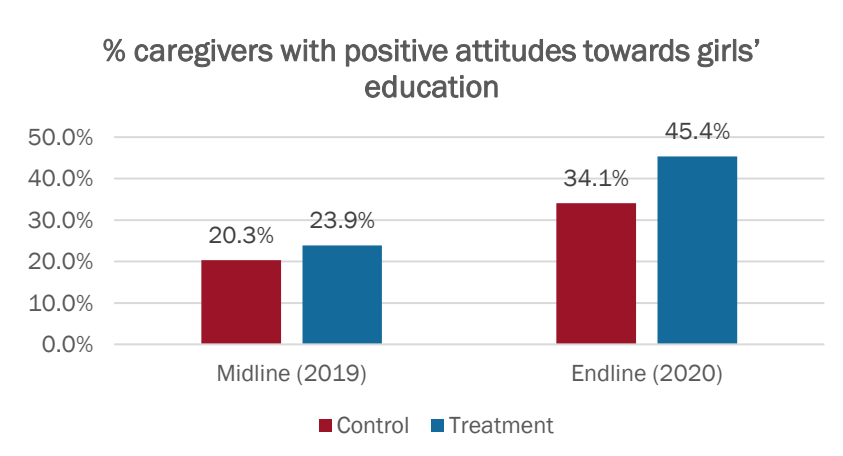
According to linear regression results using a DiD approach to measure the project's impact on time spent on chores, the project did not have a statistically significant impact on time spent¹³³.

Qualitative evidence suggests this may be due to the role of COVID-19 conditions on suppressing these outcomes.

Caregiver/parental attitudes is measured through a 10-item attitudinal scale¹³⁴ at each period.

By 2020, 45.4% of caregivers in the treatment group have a positive attitude towards girls' education compared to 23.9% in 2019.

In 2020, there is a statistically significant association between having positive attitudes towards girl's education and being in the treatment group suggesting the project has influenced attitudes amongst parents and caregivers towards their daughter's education.



¹³³ $p > 0.05$

¹³⁴ 10 item scale with 4 reverse coded items: HHS - Q314 "If necessary, parents should be able to keep their children at home during school hours to work or help in the household."

HHS - Q315 "A family has a son and a daughter but can only afford to send one of them to school. It would make more sense for them to send their son to school."

HHS - Q320 "It is more important for a woman to be a good wife and mother than to be educated."

HHS - Q322 "In general, a boy is more likely to use his education when he leaves school than a girl."

HHS - Q316 "Even when funds are limited it is worth investing in [GIRL]'s education"

HHS - Q317 "A girl is just as likely to use her education as a boy"

HHS - Q318 "Even if my daughter got married I would still encourage her to continue with her education."

HHS - Q319 "The more education a girl has the more she will be able to find good work"

HHS - Q321 "Educated women are better mothers and have healthier children"

HHS - Q323 "The education of girls is just as important as the education of boys."

Caregivers in the treatment group do not have positive attitudes towards girls' education if their daughter does not speak the language of instruction.

A chi-square test for association ($p < 0.05$) finds that girls not speaking the language of instruction in the treatment group is associated at statistically significant levels with parents not having positive attitudes towards girls' education at endline. 100% of treatment girls in the sample who did not speak the language of instruction had parents who did not view their girl's education positively.

Similarly, in the control group, a statistically significant association was found between the two variables, with 90% of girls not speaking the language of instruction and having parents who did not view their education positively. These findings suggest that project activities targeting this outcome were not successful in improving parental attitudes for girls who do not speak the language of instruction.

At endline, head of households in the treatment group who had no formal education or who were unemployed or not paid in cash or kind for work are more likely to have negative attitudes towards girls' education.

At statistically significant levels ($p < 0.05$), chi-square tests for association found that 80% of head of households who had no formal education and 79% of head of households who were unemployed or not paid in cash or kind for work in the treatment group, did not view their girl's education favorably.

Early marriage remains a barrier to girls' education in communities with some evidence that this has become more prevalent in both treatment and control areas.

Parents and caregivers in both treatment and control groups have largely similar views on the extent to which child marriages have become either more or less common in the past year. 17.2% in both groups report that child marriages have become more common since COVID-19. 67.3% of parents in the control group and 67.2% in the treatment group report that child marriages are less common.

20% of parents in treatment schools report that girls in their communities usually get married below the age of 18. 22% of parents in control areas also reported this to be the case.

Girls in validation sessions in Arsi chose to discuss this finding further. They reported that it has become more common for girls with "economic problems" to choose marriage as an alternative option to school and that some families consider marriage a "solution for their poverty ... like giving children for

exchange of cattle”¹³⁵. Some girls also reported that living with parents and caregivers during COVID-19 added additional pressure and created reason for some girls to want to get married so they could leave home.

Girls in validation session in South Gonder and South Wollo also reported increases in early marriage. As a girl explained:

“As we observe in our community and having friend who have faced this challenge the finding is correct. Early marriage is still here and a challenge for girls. Of course, before COVID-19 it was reducing from time to time as the parents have got a clue about the Girls’ Education Challenges, but during COVID-19 the parents and the girls were hopeless...in re opening of school, so many of our friends were engaged”¹³⁶.

2.3.2 What worked and did not work to improve parent and caregiver attitudes and behaviours?

Outreach activities conducted through family hubs and community volunteer networks to support girls and keep them in education, was successful in improving parent attitudes towards girls’ education.

Having a girl in a household, or a household who have been directly approached by CHADET in the last 3 years and spoken to about attendance and enrolment in school was a statistically significant predictor of improved parental attitudes between Midline and Endline. This suggests that the outreach activities are successful in supporting parents to improve their attitudes towards girls’ education.

Family hubs, in particular, which were reinstated in Q8 of GEC-T at the request of direct feedback from girls, have provided a space where girls can raise and discuss issues with parents and community leaders such as early marriage and domestic labor which may create barriers to attending school.

Being approached by CHADET in the last 3 years and spoken to about attendance and enrolment may have influenced the number of hours girls spent doing chores each day at endline.

Findings suggest that outreach activities involving community workers speaking to caregivers about the domestic chore burden at home and on attendance and enrolment, may have had an influence¹³⁷ on the mean hours girls spent on household chores at endline in comparison to households that weren’t

¹³⁵ Validation session with girls in Arsi, April, 2021

¹³⁶ Validation session with girls in South Gonder, April 2021

¹³⁷ $t(790)=-2.34, p=0.02$

approached by CHADET. Girls whose caregivers had spoken to someone from CHADET spent on average 3.37 hours on chores at endline in comparison to 3.68 hours spent on chores by girls whose caregivers were not approached by CHADET.

2.4 Teaching & Learning

2.4.1 What impact did the project have on teaching quality and learning outcomes?

The project aimed to improve the quality of teaching through the provision of teacher professional development, coaching, and mentorship and support to education officials at the Woreda level.

Specifically, in the GEC-T phase the project:

- Trained 484 teachers in: i) improved pedagogy, ii) gender-sensitive teaching, iii) child safeguarding, iv) improved literacy and numeracy instruction, and v) inclusive teaching strategies
- Trained 336 school stakeholders including school staff, Community Care Coalition and school board members, on safeguarding, case management and school leadership
- Established 8 ICT Centres across secondary schools
- Established 10 teacher resource centres at Woreda Education offices
- Set up 109 Communities of Practices (COPs) where teachers can discuss improved instructional practices
- Supported the drafting and production of 49 grade level distance learning material in 3 subjects (mathematics, English and local language) in response to COVID-19.
- Produced teaching resources on teaching children with different impairments
- Trained 492 teachers (Arsi 89, SW 138, SG 265) on supporting children with learning differences and psycho-social support to children
- Provided awards to girls based on their achievement in school
- Provided girls with exercise books, scholastic materials, and school uniforms necessary to participate in school
- Established 74 Homework Tutorial Clubs facilitated by trained tutors to provide supplementary learning to marginalized girls

Lesson observations conducted at Midline indicate that the project has supported teachers to adopt improved preparation, assessment, and pedagogy practice in their lessons.

At Midline:

- 46.8% of lessons in the treatment group demonstrated improved preparation¹³⁸, compared to 28.8% of lessons in the control group.
- 16.0% of lessons in the treatment group demonstrated improved pedagogy¹³⁹, compared to 5.5% of lessons in the control group.
- 38.3% of lessons in the treatment group had adopted improved assessment¹⁴⁰ practices compared to 34.2% of lessons in the control group.

Due to ethical concerns, the decision was taken to not conduct lesson observations endline in consideration of the increased burden that has been placed on teachers since the onset of COVID-19 and the additional pressures observations may impose. Teachers in Ethiopia currently have a heavily increased workload due to catch up teaching, social distancing and other protocols put in place to protect children and keep schools open.

In interviews, teachers trained by the project explained how improvements in assessment, preparation, and pedagogy work to support teaching and learning, suggesting the project has played a role in changing teachers' mindsets.

Qualitative sessions at endline suggest that the project supported teachers to understand and adopt improvements in each of the three targeted domains (preparation, assessment, and pedagogy). Teachers consistently reinforced the relevance of improvements across these domains in their discussions and how these changes had influenced learning:

¹³⁸ See section 2.6 in the Midline Report; lesson observations measured extent to which: The teacher uses an appropriate and well written lesson plan, lesson objectives are clearly displayed at the start of the lesson, Lesson objectives are clearly explained to the students at the start of the lesson, the teacher links the lesson to the previous lesson

¹³⁹ See midline report, section 2.6; The measurement of pedagogy was based on the extent to which: the teacher helps students solve problems, students work collaboratively (in pairs /groups), students spend more time on learning tasks than listening to the teacher, the teacher names individual students to answer questions; not just those who raise their hands, the teacher asks 'open questions' to test the students' understanding; not just repeating, or saying yes/no, the teacher supports less advanced students and ensures that the more advanced students have work that 'stretches' them, the teacher includes all students in the lesson (girls and boys equally), the teacher actively tried to involve students who were not participating, students ask the teacher questions, the teacher uses language that is clear, simple, and appropriate to the grade/age level of the students, the teacher listens attentively to students, the teacher praises the students, the teacher checks that the students are making progress against the lesson objectives at least twice throughout the lesson

¹⁴⁰ See midline report; lesson observations assessment practices measured through extent to which: the teacher gives formative feedback on the homework / tasks completed by the girls, the tutor/teacher gives a summary at the end of the tutorial, the tutor/teacher checks learning has advanced since previous tutorial, The tutor/teacher sets individual learning targets for the next tutorial

“If I give some group work, I will assess how many students fail and how many students pass then if many students fail on one question I will prepare a lesson plan in the future to revise the topic again”.

“I have now better knowledge on how to approach students, how to understand students’ problems and how to differentiate based on their understanding ... some students can understand by explanation and some students can understand by touching things ... identifying the right approach for each kind of students helps the teaching process”.

Teachers trained by the project emphasized the importance of continuous assessment, in supporting them to make decisions about how to differentiate to the levels of different groups of children.

Teachers commented:

“In a certain class there are different students but as per the training, I will prepare the plan to fulfill the need of every student. The point is about choosing the method of teaching that fits with the need of students.”

“We use continuous assessment, for example when we say assessment it can be verbal assessment, or it can be project works. You can choose one and can assess students. From the training what we get is before the training we think that assessment is exam but now after the training assessments can be by observation, or by group work, or by different methods that CHADET gives as manual.”

“When I come to school I ask students what the last topic was about and then I assess them or in the middle of class. To me the change is that before the training I ask 5 or 6 students; if they answer I feel they understood me. However, after the training I have to do an assessment that is inclusive of the three categories of students. Secondly what I get from the training is I give an assessment to the class and if one question is not answered by many students I will revise the topic again.”

Teacher survey results also suggest the project has continued to improve teachers’ attitudes towards the adoption of improved preparation, pedagogy, and assessment practices. [based on ML findings currently to be added after quant data collection in phase 2 for teacher survey]

Interviews with teachers suggest that the project increased their understanding of active and experiential learning, as well as strategies to promote group work and paired work, in particular.

Several teachers interviewed explained that the training particularly made them more aware of different ways to teach which involved more active approaches rather than “chalk and talk” teaching, and different ways to engage children in group and paired work. Several teachers also emphasized how this has positively contributed to student engagement.

During the validation workshop stakeholders agreed with this finding:

“Our group has agreed with the finding: and our evidence is: group and pair work have now become the usual practice of teachers. Initially, it was teacher-centered, but now it has changed to student centered. Teachers have showed a change in attitude to implement active learning methods.”¹⁴¹

“Before the project, girls were not participating. teachers were doing chalk and talk, where only the teacher speaks. But now, group and pair work came with the project.... girls participate, experiences are shared, and competence enhanced, self-confidence develops and this affects their academic performance”¹⁴².

The project had a statistically significant impact on numeracy outcomes between 2018 and 2019 according to the quasi-experimental impact model.

The project had a statistically significant impact on numeracy outcomes between baseline and midline, based on the cross-sectional difference-in-difference model¹⁴³. Girls in the treatment group on average improved their numeracy by 12.5% more than girls in the control group.

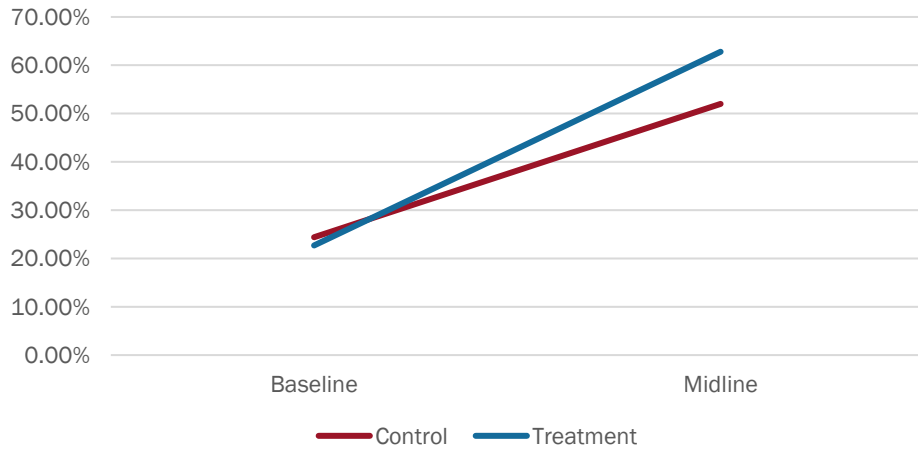
Girls in the treatment group outperformed girls in the control group in all grades with regards to meeting expected numeracy curriculum competencies, according to numeracy assessment results at Midline, further evidencing the contribution of the project to numeracy outcomes.

¹⁴¹ Validation workshop

¹⁴² *ibid*

¹⁴³ $p < 0.05$; $\text{Beta} = 12.481$; $R^2 = 0.323$; Numeracy at Midline was measured through the single subtask that overlapped between periods the Advanced written problems subtask

Average change in numeracy between baseline and midline by evaluation group (written task %)

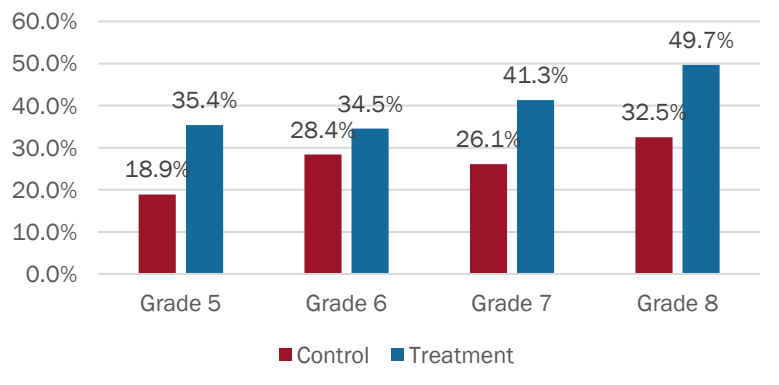


The greatest change in numeracy outcomes between baseline and midline was experienced by girls who were in grade 8 at the start of the project.

These girls would have been exposed to the greatest number of years of the project’s first phase, likely exhibiting sustained benefits from prolonged exposure to the intervention.

Girls in this grade level experienced improvements which exceeded improvements experienced by their peers in control between baseline and midline by an average of 17.2%. Scores in the written task were used to understand changes in numeracy between these periods¹⁴⁴.

Mean Change in Numeracy between Baseline and Midline by Original Cohort Membership and Evaluation Group



¹⁴⁴ Due to limitations at baseline the written task was the only task that was comparable between evaluation periods. However this task contains a range of advanced written problems representing numeracy sub-domains such as algebra, geometry, and equations with unknowns.

The project had a statistically significant impact on girls' NEAEA Grade 8 (Primary School Certificate) exam results between 2017 and 2019.

The evaluation gathered Grade 8 exam results for a representative sample of girls in each of the last three exam cycles (Ethiopian Calendar 2010, 2011, 2012) from schools in project and non-project areas across all three years. This included a sample of 1,511 girls in non-project areas (control) and 1,514 girls in project areas (treatment)¹⁴⁵.

Grade 8 examinations are managed through the National Educational Assessment and Examinations Agency and are standardized examinations covering the following subject areas: local language (Amharic or Afaan Oromo), English, mathematics, physics, civics and social studies, biology and chemistry.

A cross-sectional impact model finds that the interaction variable¹⁴⁶ was a statistically significant predictor of overall exam results, when looking at the period between the first and last exam cycle gathered (2017 and 2019)¹⁴⁷.

Mean results indicate that whilst on average exam results in control schools decreased between 2017 and 2019¹⁴⁸, exam results in project schools stayed largely the same, due to the project's support¹⁴⁹.

When we examine each year independently, the project did not have a visible impact on overall exam results between 2017 and 2018.

However, between the last two exam cycles (2019 and 2019), a cross-sectional linear regression model finds the interaction variable to be a statistically significant predictor of overall scores and that girls in project areas had an average change of 5.7% in exam score greater than the change experienced in non-project areas.

The evaluation also gathered data for three subject areas of the exam: local language, English, and mathematics. Examining results for each of these subjects suggest that the project's impact on overall exam results was driven largely by impact on the English literacy section of the exam.

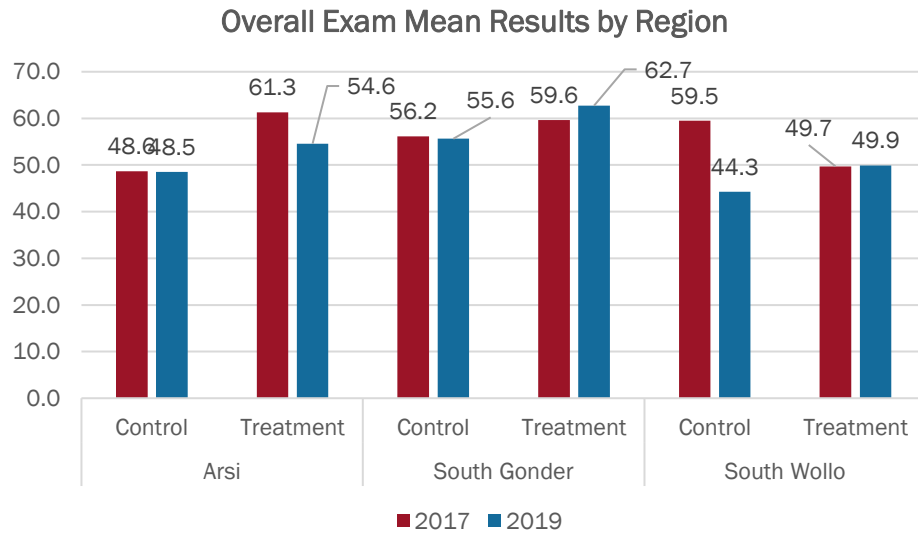
¹⁴⁵ 2010 exam data: 375 G8 girls in control and 377 in treatment; 2011 exam data: 378 G8 girls in control and 378 in treatment; 2012 exam data: 380 G8 girls in control and 381 in treatment; aggregate populations compared between periods using cross-sectional approach with time, treatment, and interaction predicting outcomes levels

¹⁴⁶ The product of time and treatment status as per cross-sectional impact methodology to isolate effects of the intervention

¹⁴⁷ $p < 0.05$; Beta=6.072; R square = 0.058

¹⁴⁸ From a mean of 56% in non-project areas to a mean of 50% in non-project areas

¹⁴⁹¹⁴⁹ Mean scores for schools in project areas held steady at 56% between the two years



Examining results for each of these subjects suggest that the project’s impact on overall exam results was largely driven by impact on the English literacy section of the exam. Between 2018 and 2019 for example, whilst girls in grade 8 non-project schools on average decreased their English literacy exam scores by an average of 6%, there were no statistically significant differences in girls’ grade 8 English literacy exam results between 2018 and 2019.

There is evidence that the project had a statistically significant impact on English literacy outcomes.

Impact on literacy could not be measured between baseline and midline due to the absence of reliable literacy data at baseline. At midline, EGRA English learning assessments were conducted for the cohort of girls in both treatment and control schools. Learning data was not conducted at Endline due to ethical concerns outlined previously in this chapter.

In the absence of a data source that could compare existing English literacy data to Midline, the study focused on collected girls’ semester grades in English for the three semesters prior to schools closing in March 2020. These internal assessments are not standardised across grades and schools but, in the case, where they have been taken in the same academic year, are provided by the same teacher.

A review of grade level data between Semester 1 and Semester of the Academic Year 2018-2019, finds that the project had a statistically significant impact on English literacy, as measured through semester grades.

The impact model, following a difference-in-difference design, i.e. using treatment status to predict the first difference, was significant and indicated

that the project contributed to a 2% increase in English semester grades¹⁵⁰. This estimated contribution represents a change over and above changes which would have been experienced without the project¹⁵¹.

Though this measure of academic achievement (grades) is non-standardized and based on internal assessments, this finding gives an *indication* that the project has *contributed* to improved English literacy outcomes, in the absence of other data sources.

It should be noted that the period examined in the analysis, a single semester, is a small window within the wider 4-year intervention. The fact that a statistically significant improvement can be attributed to the project within this window suggests the project may have had impact beyond this single semester.

Furthermore, a far higher proportion of girls in the treatment group improved their outcomes within this period than in the control group, and where these changes happened, the magnitude of these improvements was greater than the magnitude of improvements experienced by girls in the control group.

64.8% of girls in the treatment group who increased their English literacy levels between semester 1 and 2, increased these levels by 5% or more compared to only 46.8% in control.

At midline, in 2019, the study measured English literacy through the Early Grade Reading Assessment and the Secondary Grade Reading Assessment. However, English literacy was not measured by the previous evaluator at baseline thus making comparisons impossible between baseline and midline.

In 2019 girls in the treatment group had higher levels of English literacy than girls the control group at statistically significant levels, signalling that the project may have played a role in improving English literacy levels between baseline and midline.

At Midline, in 2019, 17.3% of girls in the control group and 13.4% in the treatment group were non-readers in English literacy, reading at a fluency rate of 0 to 5 words per minute.

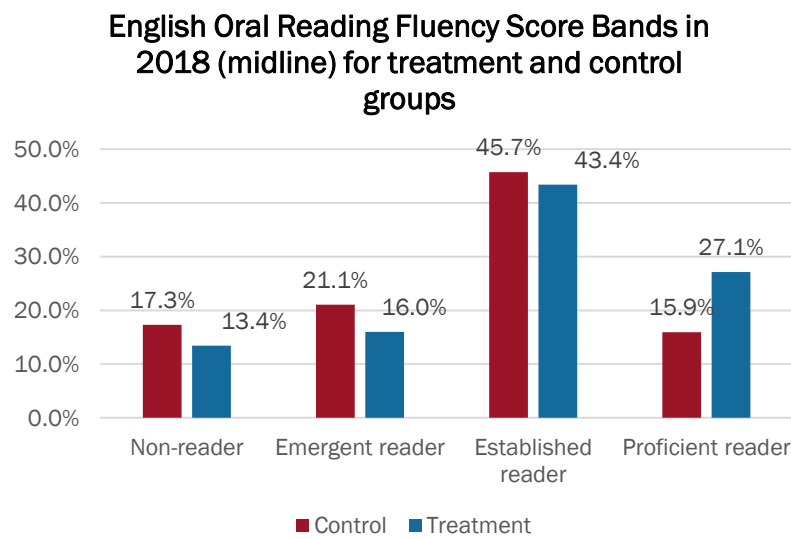
A chi-square test finds that being a non-reader is associated at statistically significant levels with being in the control group¹⁵². This suggests that the project was successful in supporting improvements in English literacy levels between baseline and midline. However, due to the limitations in reliable data

¹⁵⁰ (p<0.05; Beta=2.095; r square=0.01)

¹⁵¹ This is based on the interpretation of the Beta value in the linear regression model.

¹⁵² p<0.05 – see Midline Report for additional details

on literacy levels at baseline, and a different testing mechanism at endline, this finding is inconclusive.



Non-readers in English literacy were found in all target grade levels at Midline (Grade 5-10), suggesting subsequent interventions should aim to support early literacy acquisition in English across a wide range of age groups and grade levels. There is little quantitative evidence that the project had a significant detectable impact on Amharic or Afaan-Oromo literacy outcomes, although interviews with girls and teachers suggest interventions have contributed to improvements.

Between baseline and midline local language literacy could not be assessed due to the absence of reliable local language literacy data at baseline.

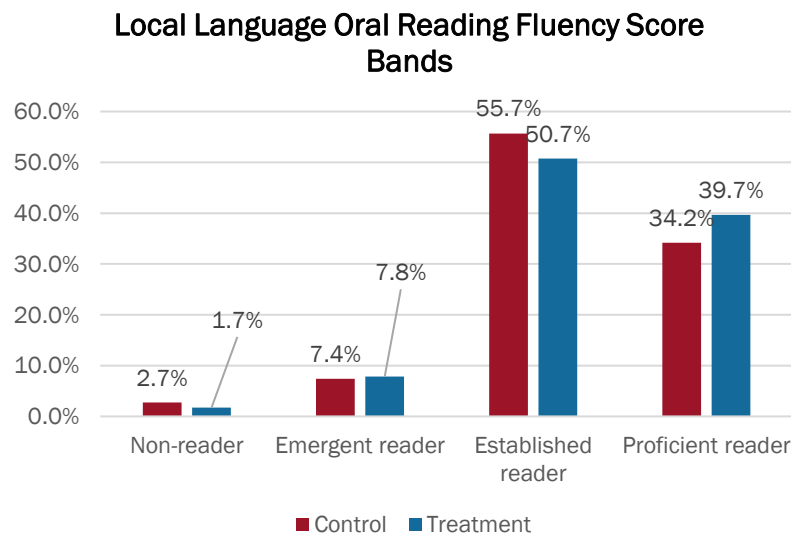
At endline, the evaluation followed a similar approach that was used for English literacy, examining local language semester grade (internal) assessment, improvements for the tracked cohort. However, treatment status was not a statistically significant predictor of improvements as measured through grades.

In 2019 girls in the treatment group had higher levels of Local Language literacy than girls in the control group at statistically significant levels, signalling that the project may have played a role in improving local language literacy levels between baseline and midline (2018-2019).

At midline, based on the oral reading fluency subtask, 1.7% of girls in the treatment group and 2.7% of girls in the control group are non-readers in local language literacy, reading at an average of 0 to 5 words per minute.

Non-readers in local language literacy were found in all target grade levels at Midline (Grade 5-10), suggesting subsequent interventions should aim to

support early literacy acquisition in mother tongue across a wide range of age groups and grade levels.



A logistic regression finds that the treatment group is a statistically significant predictor of a girls' proficiency level in local language oral reading fluency, indicating that girls in the treatment group are more likely to have higher levels of local language oral reading fluency at Midline.

This suggests that the project supports girls to improve their oral reading fluency and may have had an impact on learning in local language between baseline and midline. However, due to limitations with reliable data at baseline, and a different testing mechanism at endline, we are not able to conclude improvements linked to both evaluation points.

2.4.2 What worked and did not work to improve teaching quality and learning outcomes?

To support learning outcomes, the project trained tutors who deliver homework tutorials to marginalized girls in improved teaching practices. Tutorials provide the opportunity for English, mathematics and local language curriculum content to be reinforced and for girls to work in groups and ask questions in small group settings to teachers.

Most male and female teachers sampled through the teacher survey¹⁵³ report that the training was practical and easy to apply in their lessons.

¹⁵³ Non-representative sample of trained teachers provides an indication of what may be true in the wider population but cannot be generalized to the wider trained teacher population

A higher proportion of female teachers than male teachers report that the training they receive was practical and easy to apply in their lessons (100% compared to 90.6%).

A large proportion of teachers sampled through the teacher survey report that the time given for training was not enough for the topics covered (66.7% of female teachers and 59.4% of male teachers).

In qualitative interviews teachers often commented that they would have needed more time for the training to fully understand all of the topics that were covered. In particular teachers mentioned that they would have liked additional training on how to support children with disabilities. While this was covered in project training, most of the training curriculum was focused on the mathematics, English literacy, and local language literacy content, planning, assessment, and improved pedagogical practices.

The project should consider whether programming in the future should allocate additional time for training with teachers given that a majority of female teachers and a large proportion of male teachers feel the current amount of time was insufficient.

A higher proportion of female teachers report applying what they learned from the training in the classroom, although a high proportion of male teachers also report having done so.

97% of female teachers report that they apply what was learned in training in their classes, while 87.5% of male teachers report doing so. Additional monitoring activities in the future should aim to understand these differences between male and female teachers and if the training is well targeted to both groups.

Male teachers trained by the project were more likely than female teachers trained by the project to assess children at the start of a unit rather in the middle of a unit, the opposite was found for female teachers trained by the project.

More male teachers report pre-assessing student levels at the start of a new unit (81.8%) than in the middle of a unit (63%). More female teachers in contrast report assessing students in the middle of a unit (79%) than before starting a new unit (58%). These associations are significant at statistically significant levels according to Chi-square tests.

There are benefits to assessing students at different points. The project should consider emphasizing this in future programming for both male and female teachers. Lesson observations conducted in the future should also seek to

further investigate differences in the assessment practices used between male and female teachers given this observation.

Teacher survey data suggests¹⁵⁴ that 44.6% of teachers trained by the project are members of communities of practice (largely similar between male and female teachers)

Of members of communities of practice, 86.2% report having received direct coaching through the CoPs. 44% of those who received direct coaching rated the quality of this coaching as very good, 36% as good, 20% as average and 0% as either poor or very poor).

46.2% of girls supported by the project were members of homework tutorials (est. 7400 marginalized girls) 99.4% of tutorial members at endline report that they find these useful.

Participating in homework tutorials supports English literacy improvements and contributed to project impact on English literacy.

Participating in homework tutorials was a statistically significant predictor of having improved girls' English literacy between semester 1 and semester 2 of Academic Year 18-19, the period in which the evaluation has been able to isolate project impact on English literacy¹⁵⁵.

Similar findings at Midline further suggest homework tutorials support learning in local language literacy, numeracy, and English literacy.

Linear modeling at midline found that participating in homework tutorials predicts higher levels of English oral reading fluency, local language oral reading fluency, and numeracy at statistically significant levels¹⁵⁶.

Attending school more often leads to higher grades in English and mathematics, according to linear modelling.

No relationships were identified between attendance outcomes and grades in local languages, likely due to the fact that these are no longer used as the language of instruction in higher grade levels for most girls in the sample¹⁵⁷ and so attending school more often is unlikely to influence existing levels.

¹⁵⁴

¹⁵⁵ (p<0.05; Beta=2.065; r square=0.01).

¹⁵⁶ See midline report for regression analyses

¹⁵⁷ Oromiya is the language of instruction until Grade 8 for girls in Oromia but these girls are a minority of girls' included in the sample

Average monthly attendance levels¹⁵⁸ in 2018 and 2019 are statistically significant predictors of girls' grades in mathematics¹⁵⁹ and English¹⁶⁰, indicating that the more a girl attends school, the better her grades in these subjects.

Average monthly attendance in 2018 and 2019 are statistically significant predictors of having improved one's grades in English and in mathematics in the Academic Year 2018-2019¹⁶¹.

Supporting girls to feel comfortable participating in class, supporting their self-esteem, and supporting their self-efficacy all play a role in supporting girls learning outcomes.

Supporting girls to feel capable and comfortable participating in class is a statistically significant predictor of English oral reading fluency and numeracy at Midline ($p < 0.05$)¹⁶².

Several activities were found to promote girls' perceived capacity to participate in class including Girls' Clubs and Homework Tutorials. These findings also suggest that the training teachers have received to increase girl engagement and involvement in class has contributed to improvements in learning, and may have contributed to the project's impact on numeracy.

Self-esteem¹⁶³ at midline was a statistically significant predictor of numeracy, English oral reading fluency, English aggregate score, local language oral reading fluency, and local language aggregate score. This suggest that self-esteem and learning are likely mutually reinforcing. .

Academic self-efficacy¹⁶⁴, at Midline was a statistically significant predictor of English aggregate score, English oral reading fluency, local language oral reading fluency, and local language aggregate score at midline. This suggests that supporting girls to feel more confident completing academic tasks supports their learning in literacy and numeracy.

Several home environments factors relating to caregiver engagement were shown to support learning.

¹⁵⁸ As measured through standard attendance month without seasonal or other effects

¹⁶¹ English ($p < 0.05$; Beta: 2.457; r square 0.007); Mathematics ($p < 0.05$; Beta=5.125; r square = 0.027)

¹⁶² See midline report for regression analyses

¹⁶³ Self-esteem at Midline and Endline was measured through the Rosenberg self-esteem scale, a widely validated measure of self-esteem comprised of 10 items

¹⁶⁴ Academic self-efficacy was measured through a 3 item scale asking girls the extent to which they agreed or disagreed with the following statements: "I feel confident in my ability to learn; I feel capable of learning the material in school.; I am able to achieve my goals in school

Parental attitudes towards girls' education and increased levels of engagement are a statistically significant predictor of local language oral reading fluency and local language aggregate score at Midline.

When parents are supportive of girls' education, this supports higher levels of local language literacy. Additionally, having an adult at home to help a girl with homework is a statistically significant predictor of English aggregate score and numeracy at midline and having an adult at home to ask a child about what they do in school is a statistically significant predictor of local language aggregate score, and local language oral reading fluency.

These findings suggest that the project's activities targeting parental attitudes and parental engagement have likely contributed to achievements in learning.

2.5 Transition

2.5.1 What impact did the project have on transition?

There was an overall increase in transition rates in both treatment and control between midline and endline.

This is due to the automatic transition policy that was applied in all schools nationally, whereby all children automatically transitioned to the next grade level. This policy was put in place after schools re-opened so as not to hold a significant proportion of girls back due to school closures caused by COVID-19.

All girls in control and treatment groups, apart from grade 8 and grade 12, automatically transitioned through to the next year. Transition for grade 8 and 12 girls was determined by national examination results, which had not yet been released at the time of data collection. Therefore, data collected from girls in grade 8 and 12 have been removed from endline results to more accurately reflect transition rates.

Table 1. Overall Transition Rates

Evaluation Period		Control		Treatment	
		N	%	n	%
Endline	<i>Unsucc.</i>	2	0.5%	0	0%
	<i>Succ.</i>	406	99.5%	410	100%
Midline	<i>Unsucc.</i>	19	4.7%	16	3.9%
	<i>Succ.</i>	385	95.3%	392	96.1%
Baseline	<i>Unsucc.</i>	10	2.5%	4	1.0%
	<i>Succ.</i>	398	97.5%	406	99.0%

Between all periods, being supported by the project, treatment status, does not alter the odds of having successfully transitioned or not, at statistically significant levels, signaling that the project did not have a detectable impact on transition outcomes.

While the treatment group progressed slightly above control, these differences are not significant at statistical levels. This is likely because the transition rate in the control group was also high to begin with limiting the potential for differences to be identified.

Binary logistic regressions found that the odds of being classified as a successful transition is the same for both treatment and control cases $B = -0.088$ (0.154), Wald = 0.328, C.I. (95) = (0.676, 1.239), $p=0.915$. This suggests that transition rates between midline and endline are quite similar.

The highest rates of transition in both evaluation groups at endline are from Primary School to Secondary School.

The treatment group fared better with transitions to secondary school with 100% of all girls successfully transitioning compared to 97.9% of girls in the control group. Within school transitions have not been reported at Endline, as according to government policy, all children transitioned to the next grade level. In previous reporting, during data collection at earlier stages, some girls in Grade 8 and 9 had reported that they had not transitioned. However, this was due to the fact that data was collected during the revision period, in which girls were revising the previous year of school, as due to COVID-19 closures they had not been in school for most of that year, despite being allowed to automatically transition.

Table 2. Success Rates by Transition Types at Midline and Endline (Treatment and Control)

Transition Types ¹⁶⁵			Control		Treatment	
			n	%	n	%
Endline	All In-School	Unsucc.	2	0.5%	0	0%
		Succ.	406	99.5%	410	100%
	Within Primary School	Unsucc.	0	0.0%	0	0.0%
		Succ.	0	0.0%	38	100.0%
	To Secondary School	Unsucc.	2	2.1%	0	0.0%
		Succ.	92	97.9%	80	100.0%
	Within Secondary School (automatic transition COVID-19 Policy; N/A)	Unsucc.	N/A	N/A	N/A	N/A
		Succ.	N/A	N/A	N/A	N/A

¹⁶⁵165 No girls transitioned to vocational training or employment at baseline midline or endline so these transition pathways were removed from the table to support readability

Transition Types ¹⁶⁵			Control		Treatment	
			n	%	n	%
	Drop-out	Stayed	227	99.1%	264	100.0%
		Dropped-out	2	0.9%	0	0.0%
Midline	All In-School	Unsucc.	19	7.2%	15	5.0%
		Succ.	244	92.8%	283	95.0%
	Within Primary School	Unsucc.	0	0.0%	0	0.0%
		Succ.	94	100.0%	116	100.0%
	To Secondary School	Unsucc.	0	0.0%	1	0.9%
		Succ.	141	100.0%	109	99.1%
	Within Secondary School	Unsucc.	19	11.2%	15	8.2%
		Succ.	150	88.8%	167	91.8%
	Drop-out	Stayed	404	100.0%	408	100.0%
		Dropped-out	0	0.0%	0	0.0%

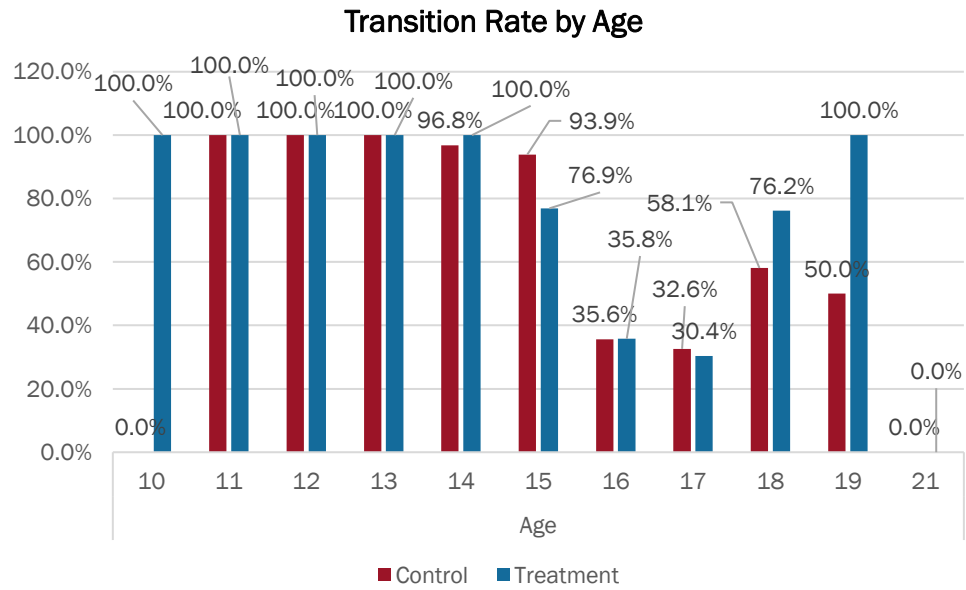
Being in the treatment group does not determine the odds of having successfully transitioned at a particular age.

No statistically significant differences were found between the control and treatment groups when disaggregating by girls' age at endline. However, successful rates of transition drop as girls get older in the control group. Transition rates for girls aged 15 and below are above 90%, while rates for girls aged 16 and above are below 60%. Transition rates in the treatment group also drop below 50% when girls are aged 16 and 17 years old but increase once more as girls get older. According to project documents, girls in this age group could most likely have repeated Grade 9.¹⁶⁶

Table 3. Transition Rates by Age

Age at Endline	Control				Treatment			
	Unsucc.		Succ.		Unsucc.		Succ.	
	n	%	n	%	n	%	n	%
10	0	0.0%	0	0.0%	0	0.0%	2	100.0%
11	0	0.0%	3	100.0%	0	0.0%	7	100.0%
12	0	0.0%	6	100.0%	0	0.0%	27	100.0%
13	0	0.0%	30	100.0%	0	0.0%	44	100.0%
14	1	3.2%	30	96.8%	0	0.0%	23	100.0%
15	2	6.1%	31	93.9%	6	23.1%	20	76.9%
16	29	64.4%	16	35.6%	34	64.2%	19	35.8%
17	31	67.4%	15	32.6%	39	69.6%	17	30.4%
18	13	41.9%	18	58.1%	5	23.8%	16	76.2%
19	2	50.0%	2	50.0%	0	0.0%	4	100.0%
21	0	0.0%	0	0.0%	1	100.0%	0	0.0%
All	78	34.1%	151	65.9%	85	32.2%	179	67.8%

¹⁶⁶ CHADET (2018) Project Documents on Transition points – under New Ethiopian Education Policy (unpublished).



Across all three evaluation periods, no girls dropped-out of school in the treatment group.

100% of all girls tracked at baseline, midline, and endline in the treatment group stayed in school. Despite school closures at endline due to COVID-19, all tracked girls returned to school.

It is important to acknowledge, that at Baseline girls were sampled from project schools by the previous evaluator and that these girls were more likely to stay in school as they were already in school. If the sample had included girls who were out of school and had re-enrolled through project support, different enrolment and transition outcomes would have been observed. One can assume that girls who are already in school, would be more likely to continue in school than those who were previously out of school. The project has supported out of school girls but these girls were not included in the original baseline sample.

In the control group, 100% of girls that were tracked stayed in school across both baseline and midline with only 0.9% of girls dropping out at endline. Differences between the control and treatment group are not significant for this period, suggesting that the project may not have been the only a contributory factor at keeping girls in school.

Table 4. Drop-Out Rate

Drop-out Rate		Control		Treatment	
		N	%	n	%
Endline	Stayed	227	99.1%	264	100.0%
	Dropped-Out	2	0.9%	0	0.0%
Midline	Stayed	404	100.0%	408	100.0%
	Dropped-Out	0	0.0%	0	0.0%

Drop-out Rate		Control		Treatment	
		N	%	n	%
Baseline	<i>Stayed</i>	404	100.0%	409	100.0%
	<i>Dropped-Out</i>	0	0.0%	0	0.0%

Being in a treatment school does not affect the chances of having to repeat a grade.

While fewer girls in the treatment group have had to repeat a grade at endline than in the control group, these differences are not significant at statistical levels. Similarly, differences between evaluation groups at midline and baseline are not significant, suggesting that the project may not have contributed towards reducing grade repetition.

Table 5. Rate of Repetition

Evaluation Period		Control		Treatment	
		n	%	n	%
Midline	<i>Passed</i>	385	95.3%	392	96.1%
	<i>Repeated</i>	19	4.7%	16	3.9%
Baseline	<i>Passed</i>	400	99.0%	406	99.3%
	<i>Repeated</i>	4	1.0%	3	0.7%

Being in a treatment school does not affect the chances of having to repeat a grade at any particular grade level.

Across grade levels, no statistically significant differences were found between the control and treatment group. However, the highest proportion of girls who had to repeat a grade were in Form 1 (Grade 9), with 73% of girls in the control group and 70% of girls in the treatment group.

High rates of repetition at Grade 9 suggest that many students find it challenging to cope with the increased academic demands of secondary school.

Students are also allowed to repeat their grade exams for free before Form 2 (Grade 10) but if a child fails their end of year examination in Grade 10 they are not allowed to repeat the grade until they have retaken the exam. This is offered outside of school through private tuition and a minimal exam fee. This could explain why rates of repetition are lower in Grade 10 in comparison to Grade 9, as girls may not have been able to pay for a retake. The project has reported that no girls approached CHADET for fees to be paid as the tuition and examination lies outside of the government school structure.

Table 6. Rate of Repetition by Grade

Grade Level	Evaluation Status Endline			
	Control		Treatment	
	n	%	n	%
Grade 4	0	0.0%	0	0.0%

Grade Level	Evaluation Status Endline			
	Control		Treatment	
	n	%	n	%
Grade 5	0	0.0%	0	0.0%
Grade 6	1	100.0%	0	0.0%
Grade 7	0	0.0%	0	0.0%
Form 1	54	73.0%	63	70.0%
Form 2	21	35.0%	22	39.3%

According to internal project monitoring data for Grade 8, transition rates remain mostly constant over periods.

An exception to this is South Wollo, which reduced by 18.15% from 2017/18 to 2019/19. However, between 2019/20, transition rates increase once more and almost return to the rate reported in 2017/18.

Transition rates in Arsi drop in 2019/20 by 7.24% from 2018/19 which is most likely due to the turmoil in the region at that time. This may have prevented girls from attending school and therefore having to repeat a grade.

Rates in South Gonder drop 3.72% from 2017/18 to 2018/19 but then returned back to almost the same rate 2019/20.

As transition rates for Grade 8 were not available at endline due to girls awaiting exam results, internal monitoring data does help explore transition rates that would otherwise not be available at this grade level. However, it is important to note that this data has not been independently verified.

Region		Unsuccessful		Successful	
		N	%	n	%
Total		358	7.1%	4716	92.9%
2019/20	Arsi	27	11.89%	200	88.11%
	South Wollo	48	6.34%	708	93.65%
	South Gonder	14	1.42%	969	98.57%
2018/19	Arsi	10	4.65%	205	95.35%
	South Wollo	163	22.67%	556	77.33%
	South Gonder	47	5.16%	864	94.84%
2017/18	Arsi	17	9.65%	159	90.34%
	South Wollo	24	4.51%	507	95.48%
	South Gonder	8	1.43%	548	98.56%

2.5.2 What worked and did not work to improve transition?

Table 7. Transition Rate by Activity and Period (treatment only)

Project Activity	Endline		Midline	
	Unsucc.	Succ.	Unsucc.	Succ.

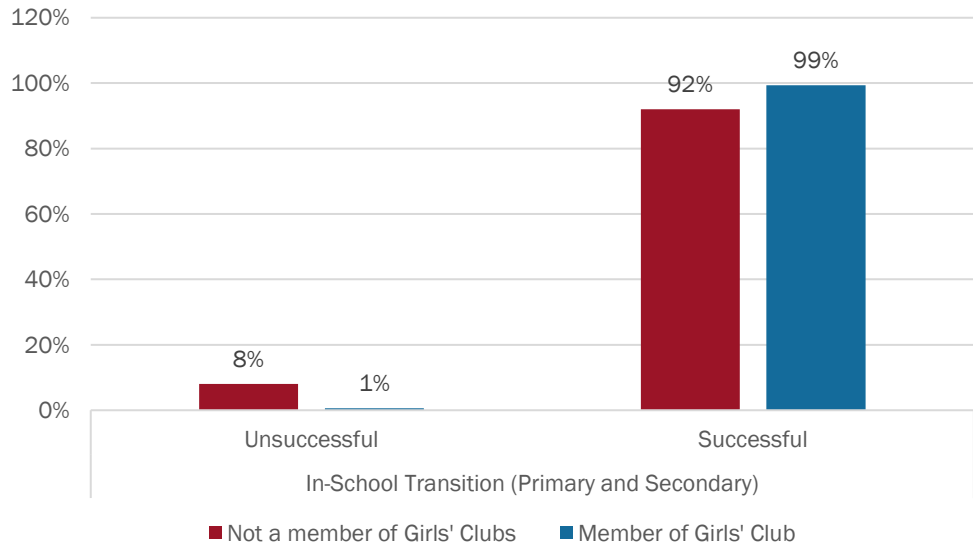
		n	%	n	%	n	%	n	%
<i>Member of HW Tutorial Club</i>	No	96	35.8%	172	64.2%	23	5.0%	440	95.0%
	Yes	67	29.8%	158	70.2%	12	3.4%	337	96.6%
<i>Made use of reading corner</i>	No	80	33.8%	157	66.2%	21	5.0%	395	95.0%
	Yes	73	33.6%	144	66.4%	12	3.4%	338	96.6%
<i>Member of girls club</i>	No	85	33.6%	168	66.4%	23	5.3%	412	94.7%
	Yes	41	31.3%	90	68.7%	2	0.9%	218	99.1%
<i>Received school uniform from CHADET</i>	No	80	32.9%	163	67.1%	20	4.7%	408	95.3%
	Yes	80	44.9%	98	55.1%	12	4.8%	237	95.2%
<i>Received leaflet about COVID-19 from CHADET</i>	No	80	33.8%	157	66.2%	19	4.6%	397	95.4%
	Yes	77	33.2%	155	66.8%	14	3.9%	345	96.1%
<i>Received financial support (scholarship/funds) from CHADET</i>	No	94	36.2%	166	63.8%	22	4.9%	428	95.1%
	Yes	9	17.6%	42	82.4%	2	2.6%	76	97.4%

Being a member of a Girls' Club is a predictor of having successfully transitioned in school at midline, according to linear modelling.

At significant levels, 99.4% of girls who were members of Girls' Clubs, successfully transitioned in-school (primary and secondary) compared to 92% who were in the control group and thus not a member of a Girls' Club. These findings suggest that being a member of a Girls' Club is a positive predictor of successfully transitioning to the next grade level.

The role of Girls' Clubs is to play an important role in developing girls' sense of belonging, self-esteem, and academic self-efficacy. The latter is especially crucial to support girls in increasing their self-perceptions of their academic capabilities. It is likely that increases in academic self-efficacy due to being a member of a Girls' Club are linked to having successfully transitioned.

In-School Transition (Primary and Secondary)



Being a member of a HW Tutorial Club and receiving a school uniform, are each a predictor of having successfully transitioned.

At significant levels ($p < 0.05$), 70.2% of girls who were a member of a HW Tutorial Club and 55.1% of girls who received a uniform from CHADET at endline, successfully transitioned.

Discussions with girls who attend homework tutorials suggest that they support girls with transitioning to the next grade level as girls have the opportunity to clarify concepts from what they learned in class and become a better student. One girl said, “I ask what I don’t understand any concept is clear for me especially mathematics and English.¹⁶⁷” Many girls see their official grade ranks increasing after attending homework tutorials, which also seems to support their likelihood of transitioning.

Qualitative findings from interviews with parents suggest that receiving a uniform was important for some girls to continue attending school, “CHADET has provided them with education materials like uniform and exercise book. This was very important really it helped us especially for a family like me who has been teaching four to five children¹⁶⁸.”

¹⁶⁷ SG19: In-depth interview with girl who improved her view of herself from HW tutorial.

¹⁶⁸ AR5: Interview with caregiver or parent of girl who is currently enrolled in school and lives far away from school.

There is a significant association with receiving financial support and successfully transitioning.

At significant levels, ($p < 0.001$). 82.4% of girls who received financial support either in the form of scholarships or funds from CHADET at endline, successfully transitioned. Providing financial support to girls to cover costs such as ~~tuition~~ fees accommodation, subsistence, and transportation was a major component of CHADET in supporting girls to transition.

These findings are supported from interviews conducted with caregivers, where financial support was deemed necessary to help keep their girls in school. For example, one caregiver said, *“I need support like covering fees for rental house. I request CHADET to do this for us. This is very important for her safety and for her education¹⁶⁹.”* Therefore, receiving financial support should help keep girls in school and increase rates of successful transitions as parents would be able to afford necessary supplies for their education.

While findings indicate that financial assistance supports girls with transition, according to project quarterly reports, the lack of financial support after the project ceases may lower transition rates considerably¹⁷⁰. Special consideration needs to be made after project closure to ensure that transition rates do not reduce.

According to results from midline, making use of a reading corner is a predictor of successfully transitioning.

At significant levels ($p < 0.05$), 96.6% of girls who reported that they made use of a reading corner reported that they successfully transitioned.

Findings suggest that reading corners provide girls with the necessary learning materials that they may not otherwise have access to in order to complete their studies and transition.

¹⁶⁹ Ibid.

¹⁷⁰ GEC-T Quarterly Project Report (QPR) Q8.

3. Relevance

3.1 How well did the project reach marginalized girls?

To understand how well the project reached marginalized girls, the study examined the composition of girls by sub-group and in relation to the intervention's intended reach and underlying assumptions about barriers to girls' education in project areas.

The project supported several key sub-groups of marginalized girls through the GEC-T phase.

During GEC1, CHADET worked with community leaders, school authorities and community workers to identify marginalized girls to participate in the project. By the end of the GECT phase the project supported 16,481 marginalized girls.

Findings at endline suggest that this included:

- **9,213 households where the head of household is unemployed (55.9%)¹⁷¹**
- **8,735 households where the head of household has no formal education (53%)**
- **5,027 female headed households (30.5%)**
- **2,506 households where the head of household cannot speak the language of instruction in schools (15.2%)**
- **1,055 households that are likely to face extreme economic hardship¹⁷² (6.4%)**
- **An estimated 445 girls with disabilities¹⁷³ (2.7%).**

¹⁷¹ This included 40% of girls households in households in Arsi, 24% of households in South Gonder, 34% of girls' households in South Wollo.

¹⁷² PPI poverty index measure Ethiopia Scorecard

¹⁷³ Defined through the Washington group short set using DISABILITY2 which includes girls with is at least one domain/question is coded SOME DIFFICULTY or A LOT OF DIFFICULTY or CANNOT DO AT ALL.

By functional difficulty domain, of these 445 estimated girls with disabilities (some multiple), 54.5% report experiencing a visual Impairment, 54.5% a hearing impairment, 18.2% a mobility impairment, 27.3% functional difficulty with communicating, 27.3% with a cognitive impairment, and 18.2% with a self-care impairment. Girls were considered to experience a functional difficulty if they reported having some difficulty or greater in the given domain.

- **33 girls who are married or living with a man as if married (0.2%)**
- **33 girls who are mothers (0.2%)**

3.2 How relevant was the project to the needs and conditions of target groups?

To understand how relevant the project was to the needs and conditions of target groups, the study reviewed the extent to which the project responded to barriers, including barriers imposed on girls' ability to continue learning due to COVID-19 and school shutdowns.

The project made timely and relevant adaptations to support girls to continue learning during school closures as a result of the onset of COVID-19.

To support girls to continue learning, the project created literacy and numeracy worksheets and homework packs with teachers. 91% of girls in the treatment group report having received worksheets from CHADET when school was closed due to COVID-19.

To support girls' psycho-social health the project also offered PSS to girls through club facilitators and tutors. 35.9% of girls in the treatment group have been selected to receive counseling from an adult over the last few months since returning to school. 100% of those who had an adult counsel them found it helpful.

The project response to COVID-19 effectively equipped girls with improved knowledge of how to prevent transmission and protect themselves and others.

88% of girls in the treatment group report having received a leaflet about COVID-19 from CHADET. 98.7% of those receiving a leaflet found it helped them increase their knowledge of COVID-19 and COVID-19 prevention strategies.

Being in the treatment group is a statistically significant predictor of COVID-19 prevention knowledge levels, suggesting the project has supported girls to know how to prevent COVID-19 transmission.

The project appropriately targeted sexual and reproductive health barriers that were found to negatively influence learning and transition for adolescent girls.

At Midline, finding it hard to access sanitary wear has a negative effect on local language literacy aggregate scores and English literacy aggregate scores, at statistically significant levels¹⁷⁴. This is likely because girls who have difficulty accessing menstrual wear would struggle to attend and learn comfortably in school during menstruation.

Similarly, not having access to someone to ask questions about SRH contributes to reduced local language literacy scores according to linear modelling at midline.

In addition, at midline the study found that not knowing a modern contraception method, including abstinence, contributes to reduced numeracy outcomes, local language literacy outcomes, and English literacy outcomes. This was a statistically significant predictor of all learning outcomes resulting in reduced levels ($p < 0.05$)¹⁷⁵.

Girls who had given birth by Midline have lower mean English literacy aggregate score, English oral reading fluency, local language literacy aggregate score and local language oral reading fluency than girls who have not given birth, at statistically significant levels. Girls who have given birth are more likely to have the burden of childcare and this may affect their engagement in school according to interviews with school stakeholders.

Chi-square tests find that girls who were pregnant at the time of the midline were less likely at statistically significant levels to experience a successful transition.

To support girls sexual and reproductive health the project has provided girls with sanitary pads, established changing areas in schools, and delivered a life skills curriculum with age-appropriate and culturally competent SRH messaging.

¹⁷⁴ ($p < 0.05$)

¹⁷⁵ Specifically, it predicted scoring 7.05% less on numeracy, 8.47% less on English literacy aggregate score, and 10.63% words per minute less on local language literacy score. This finding strongly suggests that having knowledge about how to access contraception supports learning. It is also likely that this is a proxy for wider access to SRH information, an important support for adolescents. Similar findings were found for not being able to access a condom if a girl wanted one.

Over the last year 83.1% of girls supported by the project have found it easier to access sanitary wear, compared to 53.1% of girls in non-project areas.

Chi-square tests for association are significant indicating that girls in the treatment group are more likely to report that access to sanitary wear has improved in the last year than girls in the control group.

77.4% of girls supported by the project have access to someone to ask questions to about their sexual and reproductive health compared to 62.5% of girls in non-project areas.

This represents an increase of 27.4% since midline in the treatment group (Midline value: 50%). 14.5% of girls supported by the project do not know a method of contraception compared to 30% at Midline.

The project appropriately targeted early marriage, which according to school stakeholders and community leaders has become more prevalent in communities due to the economic consequences of COVID-19 and school closures.

Quarterly reporting from the project along with interviews with community leaders, school stakeholders, girls, teachers, and project staff suggest that early marriage has become more prevalent in project communities. Quantitative evidence from the household suggests the majority of parents however believe it has become less prevalent (67.3% in the treatment group). 17.3% of reports report that it has become more common in the last year.

However, there is common agreement amongst stakeholders that early marriage remains a barrier to girls' education in the project context. Girls club facilitators report having intervened in cases where child marriages were going to take place. Facilitators emphasize that this involves approaching the parents and caregivers first and then trying to explain the barriers that early marriage cause.

The project appropriately supported girls from households facing economic hardship as these barriers had negative effects on several key outcomes.

Several findings from Midline looking at how household characteristics effect learning outcomes validate the assumption that girls in households facing economic challenges such as poverty have reduced learning outcomes:

- Living in a household where the head of household has no formal education predicts reduced local language oral reading fluency, local language aggregate score, English literacy aggregate score, and English oral reading fluency at midline. This may be because girls in these households cannot ask their head of household for help with schoolwork and have a head of household who is less likely to value education as they have not had any formal schooling.
- Living in a household where the HoH has no formal education predicts scoring 7.4 words per minute less on local language oral reading fluency, 4.2% less on local language literacy aggregate score, 6 words per minute less on English literacy aggregate score and 10.28 words per minute less on English oral reading fluency, at Midline.
- Living in a household where the head of household is unemployed predicts reduced English oral reading fluency, English literacy aggregate score, and local language oral reading fluency. This results in girls scoring 10.97 words less in local language oral reading fluency, 5.54 words less in local language English oral reading fluency, and 10.56% less in English literacy aggregate score at Midline.
- Living in a household which faces extreme economic hardship predicts reduced local language literacy, and local language aggregate score ($p < 0.05$). This results in girls scoring 8.0% less on aggregate score and scoring 11.26 words per minute less on local language oral reading fluency at Midline.

The case of one girl interviewed highlights some of the effects of economic hardship on girls' outcomes.

Mekdes lives with her mother, father and four sisters in South Gonder. Although she attends school like other children in her area, her family faces significant economic burdens. They do not have enough farming land and her mother has health problem. Mekdes also reports that her mother in particular thinks that a girls' success is measured by her marriage.

Mekdes graduated from grade 9 in 2018 but was told that her family could not afford to send her to high school because there is no high school in her area and they would need the money for rent and food. Instead, they told her that she would have to stop school and get a job to earn money for the family.

With support of the project and to respond to this she managed to start having discussions with her family where she told them that CHADET would cover scholastic materials and school uniform costs and she would work selling wood and grass on the weekends to cover the costs of rent and food, as she would be living far from home.

Mekdes now receives financial support from CHADET, attends high school and is a member of a Girls Club.

The project appropriately aimed to reduce corporal punishment in schools and the proportion of girls who report being physically punished recently by their teacher has decreased between midline and endline. However, reductions in corporal punishment cannot be solely attributable to the project based on similar improvements in non-project areas.

At midline, the evaluation found that having been physically punished by your teacher is a statistically significant predictor of reduced oral reading fluency with girls who have been punished, recently scoring 12.94 words per minute less per minute in local language ORF, on average.

Similarly, a linear regression at Midline found that having a parent use physical punishment on a girl results in girls scoring 7.26 words per minute less in English oral reading fluency on average.

At midline, of the project girls that were tracked through to endline, 8.6% reported that they had been physically punished in recent weeks by their teacher. By endline only 3.1% report this to be the case.

This suggests that increased messaging around corporal punishment has played a role in reducing this practice in schools. However, findings also indicate similar decreases were experienced in the control group. It is important to note that this may also be due to social distancing practices currently in place rather than a change in attitudes or an intentional change in discipline practices.

A higher proportion of parents and caregivers in control schools believe it is acceptable for a teacher to physically punish their children, suggesting the project may be shifting parents' attitudes: 30.1% of parents in control think this is acceptable compared to 25.4% of project girls' parents and caregivers.

The project appropriately supports girls with disabilities as findings indicate they face additional barriers to girls' education.

Quantitative findings at Midline indicate that:

- Having a mobility disability predicts reduced local language aggregate score and local language oral reading fluency. Girls with a mobility impairment score 29.9% less on local language literacy aggregate score and 39.9 words per minute less on local language oral reading fluency. Although all of these 6 girls live less than an hour to school, roads and

paths are not accessible or wheelchair friendly and this may be one reason for reduced outcomes.

- 5% of girls with a functioning difficulty were unsuccessful at transitioning, compared to 4% of their non-disabled peers. These differences are significant according to chi-square tests ($p < .001$), suggesting that girls with disabilities are less likely to experience a successful transition.

Two particular cases from interviews highlight the influence of disability on key outcomes and the role of the project in targeting these barriers.

Ababu is a 21 year old visually impaired girl in grade 12. Before secondary school she lived with her 70 year old mother and three other female household members in a rural area in Arsi.

In grade 9 Ababu moved to Dera to continue her education with the support of the project. Dera is 25km from her home. Ababu explained:

“My mother was not willing to go to Dera and continue to support me in my education. Because, there is no one who can assist me economically as well as on the way to and from school. But I have convinced her and joined grade 9, the school’s vice – principal, Ms. Firehiwot has helped me a lot. In the second semester, she has also helped me to join the CHADET project; where I have started getting the full support.

After I joined the project, I have received different support such as 1) scholastic materials, 2) medical support, 3) cash support on monthly bases, 300birr/month for 9 months, 4) walking stick, 5) braille and textbooks in grade 10, 11 and 12 made of braille 6) school uniform and above all I have received cash, worksheets in braille, sanitary pads, and soap during COVID.”

In her interview Abadu discussed specific challenges with getting to and from school without assistance, or someone accompanying her. Other girls with visual impairments interviewed in qualitative sessions raised the same concern, with some mentioning it as being a particular pertinent barrier during COVID-19 when they are not in school and otherwise having to remain at home or in their home compound.

The project appropriately supports girls in feeling safer travelling to and from school, but reductions in the proportion of girls feeling unsafe are not solely attributable to the project. Based on similar improvements in non-project areas.

At both endline and midline, chi-square tests for association ($p < 0.05$) are significant indicating that girls who live more than an hour away from school, do not feel safe while travelling to school. At endline, 28.6% of girls who travel more than one hour to school do not feel safe doing so, compared to 5.4% of

girls who do not feel safe and travel less than an hour to school. No significant associations were found between control and treatment groups in feeling unsafe and having to travel more than an hour to and from school.

Similarly, no significant associations were found between being in either the control or treatment group and not feeling safe travelling to school at either endline or midline. However, fewer girls at endline (6.3%) reported feeling unsafe travelling to and from school compared to midline (7.9%). Disaggregating by evaluation group, 7.5% of girls at midline and 5.8% of girls at endline in the treatment group reported feeling unsafe travelling to and from school. 8.3% of girls at midline and 6.8% at endline in the control group reported feeling unsafe. As rates reduced in both evaluation groups over time, this cannot be solely attributed to the project.

The most common reason girls reported feeling unsafe (28.8%) on the way to and from school was because the pathway was not safe because of others. Similar to findings at midline, the other main reasons why girls felt unsafe was because it was dark in the morning (23.1%) or dark at night (23.1%). This may explain why a higher proportion of girls who live further away feel unsafe travelling to and from school.

Evidence from qualitative sessions with girls support these findings as, when asked if harassment by others or possible abduction make it hard to travel to school, one girl said, *“Yes, it makes it hard. It is a problem. These could happen one day¹⁷⁶.”* Another girl interviewed on the same topic also said, *“sometimes girls travel alone or in pairs for long distances and could face risks like abduction, being followed by boys, harassment or abuse and even sometimes fighting and hitting¹⁷⁷.”*

While not attributed to the project or initiatives from her school, when asked what helps make her feel safe, a girl interviewed said, *“Travelling together [with two friends] makes me feel very safe.¹⁷⁸”* This suggests that CHADET’s buddy system could support girls in making them feel safe while travelling to and from school.

The project appropriately targets school safety and this remains a barrier to girls in schools.

Between midline to endline, the proportion of girls in the treatment group who reported feeling unsafe at school increased from 3.6% to 4.2%. At baseline, 1.4% of girls in the treatment group reported not feeling safe at school. Project staff have suggested that this may be cause girls’ are more aware of the risks of abduction as they get older, given the prevalence of abduction and early marriage in project areas.

¹⁷⁶ AR2: Interview with girl who is enrolled in school and lives far away.

¹⁷⁷ AR3: Interview with girl who is enrolled in school and lives far away.

¹⁷⁸ Ibid.

In the control group, the proportion of girls who reported this dropped from 4% at midline to 1.9% at endline. 3.2% of girls in the control group reported feeling unsafe at school at baseline. The only significant association found was being in the control group at baseline and not feeling safe.

Similar to midline, the most common reason given by girls in both the control and treatment groups for feeling unsafe in school was due to being bullied by other children.

At midline, 36.7% of girls in the control group and 28.6% of girls in the treatment group listed bullying as a reason for feeling unsafe. At endline, this proportion increased in the treatment group to 35.3% and decreased in the control group to 25%. No significant associations were found between feeling unsafe because of bullying and being in either evaluation group at midline or endline.

While increases in girls feeling unsafe in the treatment group are visible over time, evidence from qualitative findings suggest that girls felt safer in the last year at school. All four girls interviewed on safety all agreed that they felt safe in the semester before schools closed due to COVID-19. One girl said, *“I feel safe. We have completed the first semester in good situation¹⁷⁹.”* Another girl said, *“last year it was good, and I felt safer¹⁸⁰.”*

However, according to qualitative findings, most girls indicated that they felt unsafe at school because of boys who come into their school. One girl said, *“The school fence is very old, it has been broken and damaged everywhere. It is not safe, boys cross and get in from any direction. So, I feel unsafe going around the school fence.¹⁸¹”* Another girl said, *“I do not feel safe coming to this room. As you see it is available at the centre of the school. Many boys spend here.¹⁸²”*

The project appropriately targets bullying as this is a barrier for girls and evidence suggests the proportion of girls who are currently bullied has decreased.

9.2% of girls in the treatment group reported being bullied at Midline compared to 4.8% at Endline. Chi-square tests for association were significant at endline, indicating that bullying is associated at statistically significant levels with being the control group. The project has supported girls to be able to identify and respond to bullying and this finding indicates these activities have been successful.

¹⁷⁹ AR1: Participatory mapping girls live far away

¹⁸⁰ Ibid.

¹⁸¹ Ibid.

¹⁸² Ibid.

3.3 To what extent did the project promote gender equality and social inclusion (GESI)?

By supporting marginalized girls, the project aimed to close gaps in access and achievements across outcomes between girls and boys that are pervasive in a context that historically prioritized boys' education over that of girls and in which girls' autonomy has faced constraints.

As discussed throughout this report, there are clear gender-specific barriers faced by girls in project regions including early marriage, abduction, safety, and sexual and reproductive health barriers.

To create sustained change and respond to these conditions, the project sought to change gender norms for key groups intentionally and mainstreamed within project activities.

Parental attitudes towards girls' education improved on average between midline according to the attitudinal scale.

This finding is discussed in more detail in section 2.3. No differences were found between male and female caregivers at either period in their attitudes towards girls' education suggesting that they are largely similar.

There are no differences between treatment and control groups in the distribution of chores between girls and boys at endline suggesting the project had little influence on the distribution of chores.

At endline, 26.8% of girls in treatment schools report that the distribution of chores between boys and girls is equal compared to 26% of girls in control schools. No associations were found between equal chores and evaluation status at midline or endline.

This suggests the project played little role in influencing the distribution of household chores between girls and boys between evaluation periods.

Most girls in both treatment and control areas report that their teachers treat boys and girls equally (treatment 99%, control 98%).

Teacher attitudes towards girls' education are largely similar between male and female teachers at endline in project schools.

No statistically significant differences were found between male and female teachers with regards to their mean attitudes towards girls' education.

A higher proportion of teachers in non-project schools believe boys are always better than math than girls than in project schools (77% compared to 69%). Results were largely similar between groups when asked about reading with around half of teachers in project areas and non-project areas agreeing that girls are always better at reading and half disagreeing.

A higher proportion of girls in non-project schools than in project schools believe children with disabilities have a right to go to school, suggesting the project did not have an influence on peer attitudes towards children with disabilities. While 1.2% of girls in treatment schools report that children with disabilities do not have a right to an education, 0.7% report this in control schools.

The project appropriately targets parental attitudes towards girls' education, but findings suggest that the project did not strongly influence parental attitudes of parents of girls with disabilities towards girls' education.

According to chi-square tests for association, no significant associations were found for girls with disabilities and positive parental attitudes in the treatment group at endline.

Girls with disabilities were more likely¹⁸³ at statistically significant levels to have received a school uniform from CHADET: 72.7% of girls with disabilities in target schools compared to 60.7% of girls without disabilities in target schools.

Girls with disabilities also more likely at statistically significant levels to have attended summer transition camps (50% compared to 35%)¹⁸⁴.

Girls with disabilities however were less likely to be members of girls' clubs than girls without disabilities at statistically significant levels (GWDs: 45.5% compared to 54.1%).

Girls with disabilities were also less likely at statistically significant levels to have made use of reading corners in their school (81.8% v.s 92.9%)

¹⁸³ P<0.05 chi square

¹⁸⁴ P<0.05 chi-square

4. Value for Money

4.1 At what cost did the project deliver impact on key outcomes?

The GEC-T phase of project was delivered at an overall cost of GBP 284 per marginalized girl reached.

To answer at what cost the project was able to deliver specific outcomes, the evaluation reviewed the outcomes in which there is a high degree of probability that the project had an impact or made a significant contribution to achievements in relation to the costs of all activities directed towards these outcomes.

English literacy and numeracy were reviewed for this analysis as these were the outcomes where there is the strongest evaluation evidence that the project contributed to desired results.

While an analysis of results by particular activities, including for example Girls Clubs, the provision of school uniforms, and homework tutorials, would provide useful insights to compare value for money across specific activity lines, financial data at this level is not available from the project and would take significant human resources to disaggregate. Given that Value for Money was only integrated at Endline, financial and expenditure processes were not designed to fulfill this analysis need. However, this recommendation has been documented for future consideration.

To understand outcomes in relation to costs, the evaluation team utilized the project budget by output and assigned each output to given outcomes. M&E and administrative costs were distributed evenly across the outputs to provide an estimate of administrative costs of providing each output, as these were difficult for the project team to dis-aggregate post-hoc.

The project supported girls to improve their numeracy levels over and above improvements experienced in control at a cost of GBP 200.20 per girl.

This analysis included costs for Output 1,2, and 3 and administrative and M&E costs. It is based on the evidence from the cross-sectional impact model at midline which indicates the project had a statistically significant impact on numeracy scores between baseline and midline.

The project supported girls to improve their English literacy over and above improvements experienced in control at the cost of GBP 378.47 per girl.

This analysis included of Output 1,2, and 3 and administrative and M&E costs. It is based on the proportion of girls in the tracked cohort who improved their English literacy levels according to semester grade improvements experienced in project areas.

An estimated 52.6% of girls were able to improve their outcomes based on semester grades. However, indications from wider evidence reviewed including findings of the projects impact on national exam results, suggests a larger proportion of girls experienced improvements over the duration of the 3-year project.

5. Sustainability

5.1 How sustainable are project achievements?

To assess project sustainability, the evaluation organized a participatory session focused on sustainability with external project stakeholders during the validation workshop. Participants included teachers, headteachers, project officers, community workers, and woreda officials. The session was held after the main validation workshop, so stakeholders were aware of the main findings of all outcomes as they approached these discussions.

The GEC sustainability scorecard was adapted to only include the lowest three sustainability ratings, namely Latent (1), Emerging (2), and Becoming Established (3). In small groups participants scored the sustainability of each outcome across the community, school and system levels using the descriptors in the criteria provided. Participants then presented their justification for the score to the wider plenary group for feedback.

Scores across outcomes were averaged across regions, where a given region chose a specific outcome.

Sustainability ratings and justifications provided during the workshop are summarized in the table below

Table 8. Validation Workshop Stakeholder Sustainability Ratings by Outcome

	Community	School	System
<p>Attendance</p> <p>2-3 Emerging/ Becoming Established</p>	<p>2 (Emerging)¹⁸⁵ – Stakeholders agreed that the community has good attitudes which support girls’ attendance and behaviors have started to change at the community level with more parents supporting girls to attend school. However, barriers such as abduction, early marriage and long distances to school remain.</p>	<p>3 (Becoming Established)¹⁸⁶ – there are many actors supporting girls attendance outcomes at the school level and they have taken necessary steps to do so. Teachers for example closely monitor attendance outcomes.</p>	<p>3 (Becoming Established)¹⁸⁷ – There is strong supervision and monitoring at the regional level. There is continuous support and awareness as well as monitoring mechanisms for to monitor attendance levels from schools.</p>
<p>Learning</p> <p>2 - Emerging</p>	<p>1/2 (Latent/Emerging)¹⁸⁸ Parents and caregivers’ attitudes have changed but participants disagreed whether existing structures in the community level could sustain gains in learning after project supported activities end. Stakeholders did not identify activities at the community level that could continue to support girls learning. This could include supplementary learning activities or access to learning materials in the community. While this is largely outside of the project’s theory of change, community supports for learning are increasingly relevant particularly</p>	<p>2 (Emerging)¹⁸⁹ – Teachers have changed practices in schools and the learning environment has improved. However, stakeholders argued that the consistency of tutorials are dependent on allowances provided to teachers by CHADET which are unlikely to continue. Furthermore, stakeholders agreed that a critical mass of stakeholders (i.e. the majority of teachers in target schools) at the school haven’t adopted improved teaching practices. Stakeholders argued this was made clear by lesson observations, at Midline and based on the lack of observations at Endline it is difficult to ascertain existing levels of</p>	<p>2/3 (Emerging Becoming Established)¹⁹⁰– Stakeholders who scored 3 justified the rating as follows: at the system and policy level initiatives are continually undertaken to support learning outcomes and there are practices and processes in place to continue to support these outcomes. Those who rated a “2” reported that there was insufficient number of district officials so their influence in schools is limited despite intentions to do so.</p>

¹⁸⁵ 75% of participants in the validation workshop voted and agreed with this rating while 25% disagreed

¹⁸⁶ 58% of participants agreed with this rating and 42% disagreed

¹⁸⁷ 53% of participants agreed with this rating and 47% disagreed

¹⁸⁸ Disagreement on the rating between participants – justifications given on both sides in the two regions which discussed this outcome;

¹⁸⁹ 100% of participants agreed with this rating

¹⁹⁰ No consensus reached between the two ratings

	Community	School	System
	under conditions of school closure (also in the future)	said practices but these are unlikely to be the majority of teachers.	
Transition 2- Emerging	2 (Emerging) – Activities have supported parents, caregivers and community members to change their attitudes. However barriers still exist at the community level preventing girls from enrolling or causing them to disenroll (such as economic barriers, abduction and early marriage)	2 (Emerging) – The school environment is conducive to supporting girls to transition but there are areas where it can be improved particularly for girls as we observe differences in transition between girls and boys.	2 (Emerging) – There are organized structures from ministry to school level to create integration and coordination between schools and sectors. However, there is a need to identify best practices that can be used for replication. This requires gathering of more data.
Teaching quality 2- Emerging	1 (Latent) – Stakeholders reported that parents and caregivers’ members have a passive attitude towards teaching quality entrusting this to the schools. However, attitudes towards school in general has improved. This creates little demand at the community level for improved practices in schools therefore sustainability of these achievements at this level is latent.	2/3 (Emerging/Becoming Established)¹⁹¹ : Those who rated a rating of 3 argued that the school has a high awareness of the need to continue to support improvements in teaching quality. However, those who rated it a 2 reported that schools had not allocated sufficient funds to continue to support teaching improvements. Concerns were raised that the school could not sustain tutorials for example if external support is not provided to teachers.	2 (Emerging) – While there was commitment for teacher professional development at the district level, concerns were raised about sufficient budgeting to continue in providing teacher training at a similar level of quality that was provided by the project.
Life Skills 2- Emerging	2 (Emerging) – There is an interest and some commitment from the community to continue community activities which support these outcomes. Stakeholders believe that family hubs and safeguarding training through CC will also play a role in sustaining life skills achievements.	3 (Becoming Established)-¹⁹² School stakeholders plan to continue to run the clubs and provide scholastic materials which were shown to support these outcomes. Schools also understand the importance of these outcomes and the relationship between the provision of this belonging, self-efficacy and self-esteem.	1 (Latent)¹⁹³ : The system does not have the capacity to support the club-based activities or the provision of materials. There is little awareness on the role of life skills or their relative importance to educational outcomes. The project disagreed with this rating given that they have provided the guidance materials including the job chart and other supports.

¹⁹¹ 50% of participants rated 2 and 50% rated 1 – no consensus reached

¹⁹² 57% of participants agreed with this rating and 43% disagreed

¹⁹³ 90% of participants agreed with this rating and 10% disagreed

6. Recommendations

The following recommendations are made to inform project design and implementation in project regions in the future:

- 1. Advocate for the efficacy of teacher training on preparation, assessment, and pedagogy provided by the project in delivering key quality education outcomes.*
- 2. Advocate for the efficacy of improved pedagogy, lesson preparation, and assessment practices in delivering improved English literacy and numeracy outcomes.*
- 3. Advocate for the efficacy of bursaries, the provision of school uniforms, and other material supports provided to girls in delivering key quality education outcomes.*
- 4. Promote the importance of self-esteem, academic self-efficacy, and belonging in supporting marginalized girls' learning outcomes and wider psychosocial wellbeing, particularly with system level stakeholders. School level stakeholders acknowledge the importance of these outcomes but report that this remains unacknowledged at the system level.*
- 5. Promote the delivery of improved instruction through homework tutorials, an existing intervention activity in organized by all schools in Ethiopia. Several stakeholders in validation indicated that this will support the sustainability of project achievements in schools.*
- 6. Consider targeting girls' autonomy alongside belonging and capability: The project likely contributed to improvements in girls' belonging and capabilities. Research in psychological, particularly based on self-determination theory, has found that autonomy, capability, and belonging are important ingredients to engage in goal-directed behavior. While the project actively supported girls' belonging and capabilities, autonomy was not integrated strongly in life skills programming or in teacher pedagogical training. The*

project should consider better integrating relevant approaches including, for example elements of child-centred and inquiry-based instruction.

7. *Review project eligibility for Payment by Results from the Girls' Education Challenge.* The project had impact on numeracy outcomes at Midline and impact on national examination results between baseline and midline. Numeracy impact could be reviewed against targeting set in the outcome spreadsheet to assess eligibility for PbR based on performance.
8. *Review project assumptions on what influences transition outcomes and how programming can better support transition and address barriers influencing transition in the future.* While some project activities were found to support transition, the project did not have impact on transition outcomes suggesting that other factors influence these outcomes that are unaccounted for in the project's theory of change. Given inherent limitations in the measurement of transition, due to government policy of automatic transition, conclusions as to what these factors are at endline are limited. Project staff report that in-school policies on "detention" and holding students back may support a culture where underperforming children are more likely to be held back than identified and supported earlier on to effectively promote transition.
9. *Review project assumptions on local language literacy and how programming can better support local language literacy learning, particularly for girls in secondary schools.* The project did not have an impact on local language literacy, despite evidence that some activities support these outcomes. The project should review underlying assumptions as to how barriers can be better addressed.
10. *Review frequency and duration of teacher training to ensure it is sufficient to cover relevant topics.* Teachers in qualitative sessions and through the teacher survey indicated that the training did not provide sufficient time to cover the materials. Additionally, stakeholders in the validation workshop agreed that the majority of teachers in project schools had not adopted practices, limiting the sustainability of project achievements in teaching quality despite commitment at the school leadership and system level to support improvements. These findings suggest the project should review whether the duration and frequency of the training is

sufficient in future project design or how CoPs can be supported to provide an opportunity to re-train or cascade training to teachers. Feedback mechanisms could also be established through CoPs to ensure future training is relevant and closely addresses the needs of CoP members.

11. *Review the extent to which activities were designed to improve gender equality and norms around gender.* While the project made progress in supporting parents to improve their attitudes to girls' education, the project did not have visible influence on the distribution of chores between girls and boys, which remain similarly unequal in both project and non-project areas. Additionally, a large number of teachers in project areas believe that girls are always better at reading than boys and a smaller but concerning minority also believe that boys are always better than math than girls. These are likely proxies for wider norms and biases some teachers may hold.
12. *Consider community-based interventions aimed at targeting learning outcomes:* Based on project sustainability findings around the lack of community supports for learning, the project should consider activities that are community based to supplement supports for children who are in-school. The project spent significant resource tracing and supporting truant girls and aimed to support girls who were out of school and these children could benefit from the access to reading materials and learning at the community level.
13. *Review the extent to which Girls' Clubs are and remain inclusive of girls with disabilities.* Girls with disabilities are less likely at statistically significant levels to be members of girls' clubs. The project should review the extent to which it supported girls- with disabilities to access clubs. Furthermore, a higher proportion of girls in project areas believe girls with disabilities do not have a right to education, compared to non-project areas. While this is a small minority, it suggests additional work needs to be done on promoting inclusion amongst girls.
14. *Strengthen parent and caregiver demand for quality improvements for girls and their engagement in schools:* While the project made significant gains in parental attitudes towards girls' education, project stakeholders consulted widely agreed that parents and caregivers have a passive interest in their children's education and largely entrust schools with this responsibility.

Future programming may consider better integrating parents and caregivers into school governance structures, given the role that home environment factors play in influencing outcomes. This would further support education programming in changeable conditions such as during COVID-19 school closures, where parents and caregivers have an increased role and responsibility in supporting their children's learning.

15. *Design a robust MERL Framework and implement it consistently:* Several gaps from baseline prevented the evaluation from understanding project impact on key outcomes including literacy, self-esteem, academic self-efficacy, and belonging. Additionally, baseline indicators were not correctly operationalized and, in some cases, could not be reported on preventing comparisons in project achievements between baseline and midline.

Annexes

The following annexes submitted along with this report to the project Fund Manager:

- Annex 1 – Project Design and Interventions
- Annex 2 – Evaluation Approach, Sampling & Methodology
 - Sampling Framework
 - Inception Report
 - Endline Field Report (including information on sampling, field work, cohort tracking & attrition)
 - Midline Field Report (including information on sampling, field work, cohort tracking & attrition)
- Annex 3 – Endline Evaluation Tools
- Annex 4 – Outcome Spreadsheet & Logframe
- Annex 5 – Datasets and codebooks
- Annex 6 – External Evaluator’s Declaration
- Annex 7 – Characteristics and Barriers (added based on review to be uploaded)

Acknowledgements

There are several stakeholders which need to be thanked for their role facilitating the evaluation. Their support has ensured that the evaluation was conducted to a high standard of quality and with close fidelity to research design.

Firstly, we would like to thank all girls, parents, teachers, Woreda officials, and other stakeholders who participated in the research. Participants were often not provided with incentives to participate but still committed significant periods of time to support the study, with the understanding that it would contribute to improved educational quality in the region and in Ethiopia more generally. Each of the girls, parents and caregivers participating in the study spent around two hours completing relevant assessments at each evaluation period.

Data collection activities in-country were coordinated by James Gathogo from Health Poverty Action¹⁹⁴ and Girum Ketema. James and Girum ensured targets were met and enumerators were provided with all administrative and logistical support necessary to achieve the study's core objectives. This research could not have taken place without their drive and continued commitment to the quality of the process.

The evaluation relied on the support of a team of over 50 enumerators across the two regions of Amhara and Oromia, including the zones of South Wollo (Amhara), South Gondar (Amhara) and Arsi (Oromia). Enumerators had to closely follow field protocols, tool administration guidance, and quality control processes to sample girls and their families in remote settings.

This required an immense effort, involving contacting school stakeholders in advance and coordinating data collection activities at schools and at households, all while ensuring the study could track the same participants and that daily targets could be met. The full package of assessments took around 2 hours to administer for each case, including learning assessments, and the survey administered to girls and their head of household and caregiver.

Each zone had a qualitative research team, comprising a Qualitative Research Assistant (QRA) and one or two transcribers. QRAs in consultation with Zonal coordinators and project staff, recruited for all qualitative sessions, arranged venues, facilitated the sessions, and completed daily debriefing forms.

All audio sessions were transcribed by the transcription team, with the aim of completing the transcriptions during data collection, so the national

¹⁹⁴ <https://www.healthpovertyaction.org/>

coordination team could feedback to QRAs on progress, identify areas where the research had reached data saturation, and adapt tools accordingly. This effort required the continued commitment and drive of the QRAs and transcribers who we would also like to offer our thanks to.

The enumerator and qualitative researcher team in each zone were supported by a Zonal Coordinator who they reported to on a daily basis, and who conducted one-on-one quality assurance visits with each enumerator. The three zonal coordinators often worked after hours to ensure they could report back to the national coordination team on sampling, the progress against set targets, and changing conditions in the field which included unexpected school closures due to political tension and harvest collection due to unforeseen adverse weather conditions requiring children to participate in harvesting activities, We would therefore like to extend our thanks to the three Zonal Coordinators: Girma Fentie (South Wollo), Yohannes Kebede (South Gondar), and Mohammed Abdo (Arsi).

At the national and zonal level, the study relied strongly on the efforts of the local CHADET headquarters, offices and field staff. Field staff were incredibly helpful in supporting enumerators and QRAs to access sample sites and to regularly problem solve on-going challenges with data collection. At the headquarter zonal level we would like to acknowledge the guidance and input from Wossen Argaw (CHADET Deputy Director), Kinfe Dirba (GEC Project Manager) and Chala Legasse, (GEC Monitoring & Evaluation Co-ordinator). We would also like to acknowledge the significant support from the South Gondar team - Adebabay Fenta (Project Co-ordinator), Adem Desalegn (Deputy Project Co-ordinator), Worku Kindie (Knowledge Management Officer); ; the South Wollo team - Abraham Abebe (Project Co-ordinator) and Jemila Mengesha (Knowledge Management Officer) and the Arsi team - Fufa Beha (Project Co-ordinator) and Naol Tune (Knowledge Management Officer).

The evaluation was also strengthened by the continual support, and feedback provided by the ChildHope and CHADET teams. At ChildHope this included critical support and feedback provided by: ChildHope, UK staff: Angela Keenan (Monitoring & Evaluation Advisor), Iain Disley (GEC – Partnerships and Programmes Manager), Colin Haikin (Education Partnerships Advisor), and Jill Healey (Executive Director). The quality of the evaluation was also supported through continued consultation and partnership with the Fund Manager’s EO for this project, Iram Zahid.