Transformational Empowerment of Adolescent Marginalised Girls in Malawi – Baseline Report

April 2020

1. Cover sheet

- Name of project: Transformational Empowerment for Adolescent Marginalised Girls in Malawi (TEAM Girl Malawi)
- Names of Authors: Aimee Reeves, Anne Laesecke, Casey McHugh, Hetal Thukral, Melyssa Sibal
- Name of external evaluation firm: School-to-School International
- Date: 1 April 2020

2. Executive summary

Background

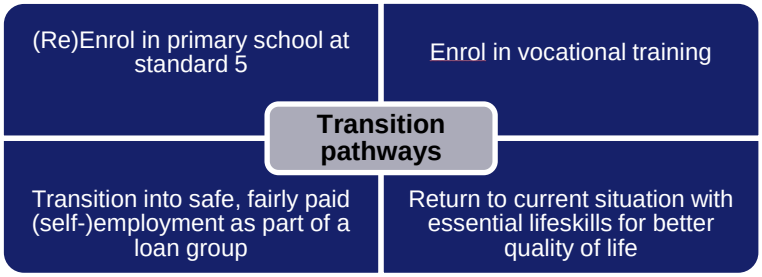
The Transformational Empowerment for Adolescent Marginalised Girls in Malawi (TEAM Girl Malawi) project is a 5-year Girls' Education Challenge (GEC) initiative funded by the United Kingdom's Department for International Development (DFID) through the Leave No Girl Behind (LNGB) funding window. TEAM Girl Malawi is implemented by Link Community Development International (Link) in collaboration with consortium partners Theatre for a Change (TfaC), Charlie Goldsmith Associates (CGA), Supreme and MicroLoan Foundation.

Seeking to improve learning and life opportunities for girls aged 10–19 who have never been to school or who dropped out of school without gaining functional literacy and numeracy skills, the project will implement activities in 4 key intervention areas:

- Community-based complementary basic education centres (CBEs)
- Girls' Clubs located in primary schools in the same communities
- Support for transition into primary school, vocational training and business training supported by micro-loans located in select communities
- Support to families, community members and government staff

The project expects to reach 6,000 direct beneficiaries in three cohorts, with 2,000 direct beneficiaries each, who will transition into one of 4 pathways (Figure 1).[1]

Figure 1: TEAM Girl Malawi transition pathways



TEAM Girl Malawi developed a theory of change (ToC) that articulates the specific barriers faced by marginalised girls in Malawi and proposes the activities, outputs and outcomes that will lead to achievement of its desired impact (Annex 15: Project's Theory of Change). The project's ToC considers multiple and intersecting barriers that prevent highly marginalised girls from accessing a quality education in Malawi, which are categorised under social marginalisation, economic marginalisation and educational marginalisation. The project's ToC proposes a set of activities that will be implemented by TEAM Girl Malawi's consortium partners to directly address these barriers. As a result of these activities, TEAM Girl Malawi anticipates the following outputs:

- CBE is high quality, inclusive and gender responsive
- Girls are empowered with sexual and reproductive health and rights (SRHR), social and emotional knowledge, attitudes and skills
- Improved leadership at national, district and local level to support the education of marginalised girls
- Marginalised girls are safe, supported, and protected
- Girls and their carers have skills to earn

[1] The TEAM Girl Malawi enrolment database included 2,009 girls and 407 boys. See Annex 11: Sampling framework.

Building on these outputs, TEAM Girl Malawi expects to observe the following intermediate outcomes:

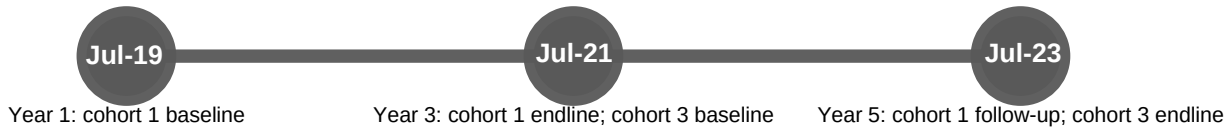
- Improved attendance at CBEs
- Improved knowledge and understanding of SRHR, self-confidence, self-esteem and wellbeing of marginalised girls
- Improved quality of education
- Improved understanding and use of support mechanisms for marginalised girls within communities
- Strengthened district and national leadership

All activities, outputs and intermediate outcomes lead to the three core outcomes of TEAM Girl Malawi–learning, transition and sustainability–which aim to improve life chances for marginalised girls in Malawi.

Approach

The evaluation of the TEAM Girl Malawi project employs a mixed-methods, longitudinal, quasi-experimental design. The evaluation utilises data from learning assessments, a package of quantitative and qualitative instruments and ongoing project monitoring tools. The variety of tools, respondents and methods of data collection allow data to be triangulated and linked across evaluation questions and indicators. Evaluation data will be collected at 3 time-points (Figure 2).

Figure 2: Project evaluation points and cohorts



Quantitative baseline data was collected in 14 CBEs and qualitative data in 4 CBEs in late June and July 2019 (**Error! Reference source not found.**).

Figure 3: Baseline sample sizes

Educational marginalisation analysis, barriers and analysis of projects' gender approach

TEAM Girl Malawi pre-identified marginalised subgroups and targeted individuals from these subgroups during enrolment (Figure 4). Barriers to learning and transition identified at baseline were explored using a mixed-methods approach (Figure 5).[2] When examining the intersection between subgroups and barriers, findings indicated that school cost was the most highly prevalent

[2] Barriers mentioned during FGDs were quantified using survey item responses.

barrier. Additionally, more than half of respondents in all subgroups, except for girls who are heads of households, experienced food insecurity or hunger as a barrier.

Figure 4: Proportion of sample by characteristic subgroup

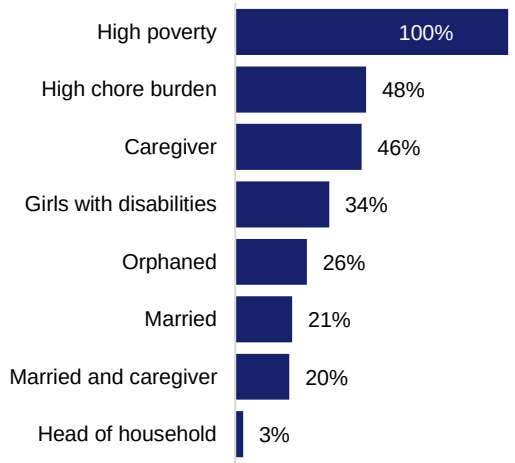
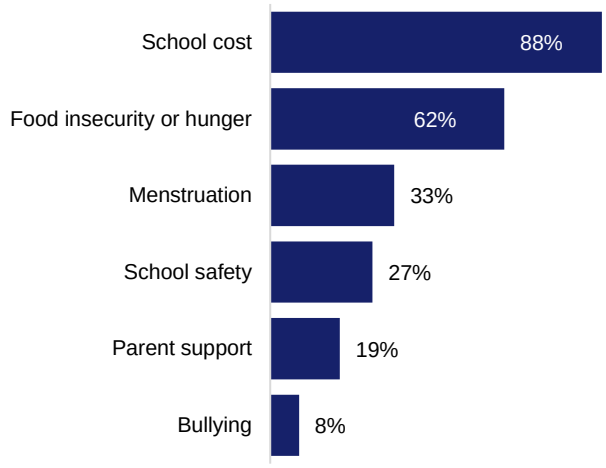


Figure 5: Proportion of sample by barrier



When examining the TEAM Girl Malawi interventions from a gender-integration perspective, the project was found to be 'gender sensitive'. The project conducted a targeted and inclusive enrolment process that reached the intended subgroups of girls. It addresses many identified barriers that restrict girls' learning and transition, and it aims to change perceptions throughout the communities where it is working. By including equally marginalised boys in CBEs, the project is aiming to reduce resentments and perceptions of favouritism while also allowing support for social-norm change and for equality.

Baseline levels

Learning[3] **—** Overall, girls' baseline literacy levels in Chichewa were very low. More than half of girls did not correctly respond to a single item on 5 of 7 literacy subtasks (Figure 6). The proportion of these 'zero scores' were high on foundational skills subtasks, including initial sound identification and syllable identification—69.84% and 64.81%—although a smaller proportion (48.41%) of girls received zero scores on the letter name identification subtask. Reading sight words also proved challenging for girls—71.43% did not read a single item on the familiar word reading subtask. Girls, however, performed better on the listening comprehension subtask, where 44.71% scored as established learners, the largest proportion on any subtask. Girls showed stronger performance in mathematics subtasks, where there is a wider distribution of performance and, in general, fewer girls who received zero scores (Figure 7). The largest proportion of girls to score as 'established learners' was on the quantity discrimination subtask— 35.98%—with similar proportions scoring as 'emergent learners' or 'non-learners' (27.25% and 27.51%, respectively). More than half of girls received zero scores on higher order addition and subtraction subtasks—

[3] Learner categories for both learning assessments are defined as non-learners who answered 0% of questions correctly, emergent learners who answered 1-40% of questions correctly, established learners who answered 41–80% of questions correctly and proficient learners who answered 81–100% of questions correctly.

53.17% on addition level 2 and 56.08% on subtraction level 2. Although 22.22% of girls received zero scores on word problems, 12.43% scored as proficient learners.

Figure 6: Proportion of girls in literacy learner categories by subtask

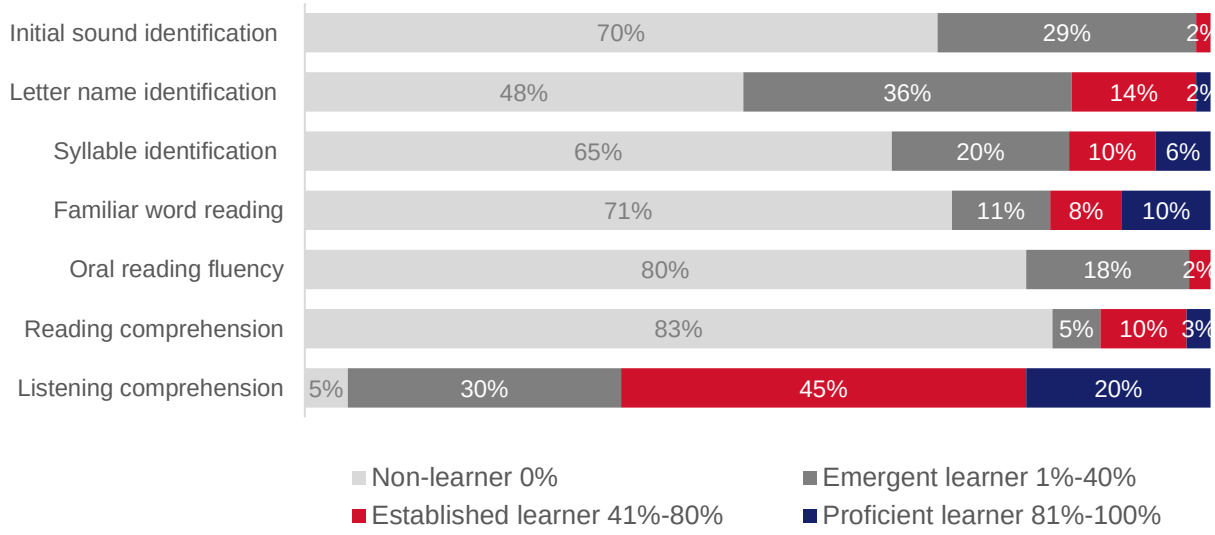
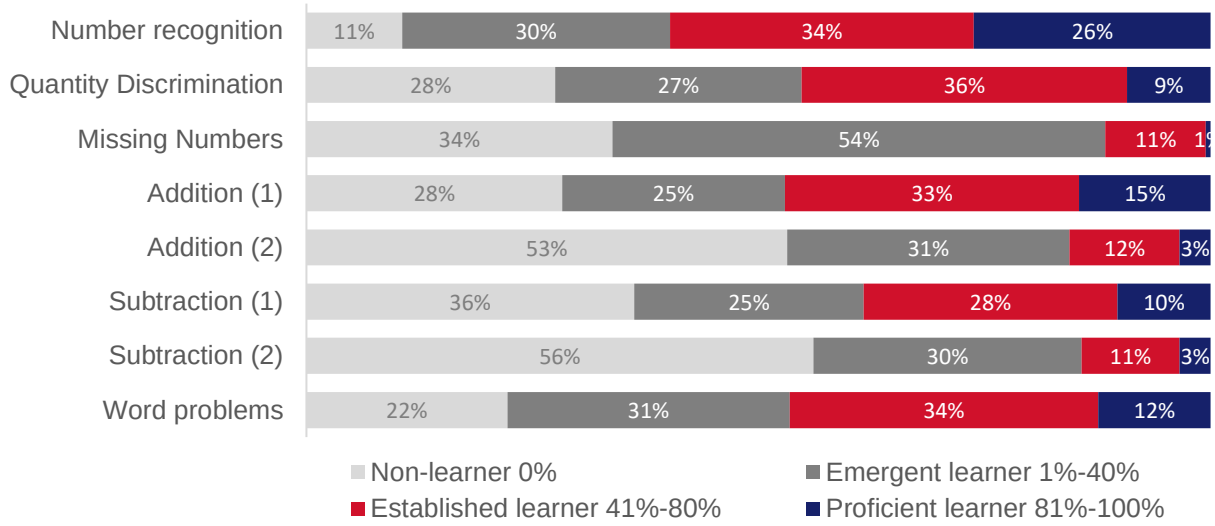


Figure 7: Proportion of girls in numeracy learner categories by subtask



The baseline also examined the ability of the learning assessments to capture growth over time. Upon examination of results, there do not appear to be ceiling effects at baseline. On most literacy subtasks, no more than 10% of girls scored as proficient learners. Although performance on mathematics subtasks was stronger, ceiling effects do not appear to be a concern.

Given these findings, the project appears to have accurately targeted girls without functional literacy and numeracy. It does appear that there is substantial room for literacy improvement during girls' 2 years of CBE.

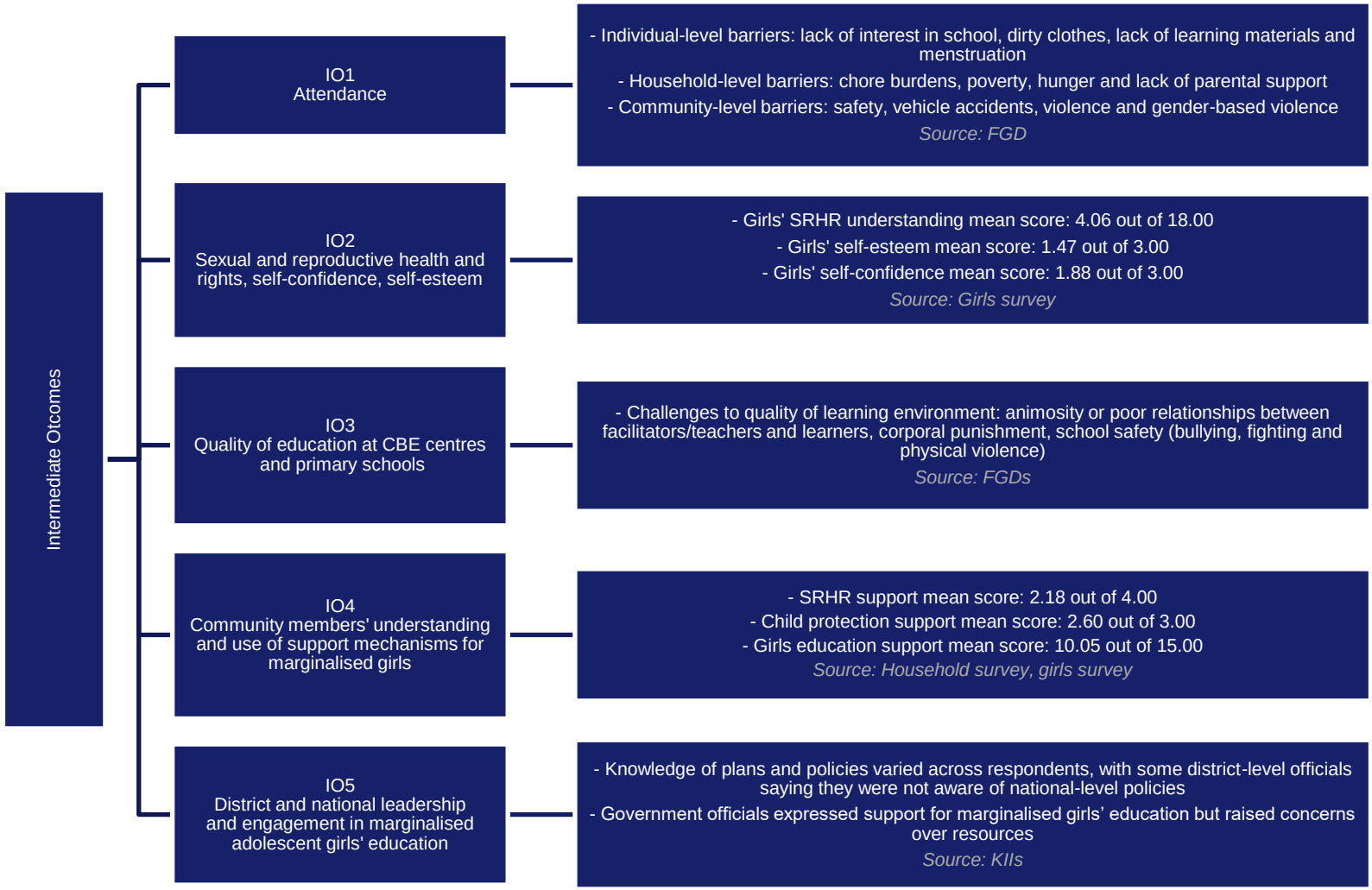
Transition — Baseline transition findings showed that 94.40% of sampled girls said they believed they would complete CBE. Of those, more than half (52.87%) reported that they hoped to go to

vocational training after completing CBE, and 39.08% said they wanted to work in a safe, fairly paid job CBE. About one in 5 girls responded that they hoped to go to primary school after completing CBE. Overall, larger proportions of girls in younger age groups reported that they hoped to return to primary school whereas larger proportions of girls in older age groups reported they hope to go to vocational training or work. These trends align with TEAM Girl Malawi transition pathways, which anticipate that girls aged 10-15 at the end of CBE will transition into the formal school system.

Sustainability — Baseline sustainability findings—presented for system, community and learning space indicators—were drawn primarily from qualitative data. The overall score on the sustainability scorecard was 1.00 out of 4.00. Evidence indicates some foundations for sustainability but also substantial room for growth. System sustainability refers to education officials' knowledge about and responsiveness to marginalised girls' educational needs. Findings varied across respondents, but all respondents named 'lack of resources' as an impediment to supporting and sustaining initiatives for marginalised girls' education. Evidence of community sustainability was more encouraging. Community leaders and other stakeholders exhibited knowledge of and willingness to engage in sustaining marginalised girls' education. Learning space evidence was limited at baseline; further research into these indicators will be collected and analysed at the next evaluation point.

Intermediate outcomes — Baseline values and key findings for intermediate outcome (IO) indicators are summarised in Figure 8. Although indicators under IO1 and IO3 were 0.00, because learning sessions had not yet begun, qualitative findings highlighted several considerations that TEAM Girl Malawi should take into account to ensure regular attendance and effective learning. Girls' sexual and reproductive health and rights (SRHR) understanding was notably low at baseline, while their self-esteem and self-confidence were relatively higher. Findings related to community support showed moderate levels of existing support, with room for growth over time.

Figure 8: Key baseline intermediate outcomes



Conclusions

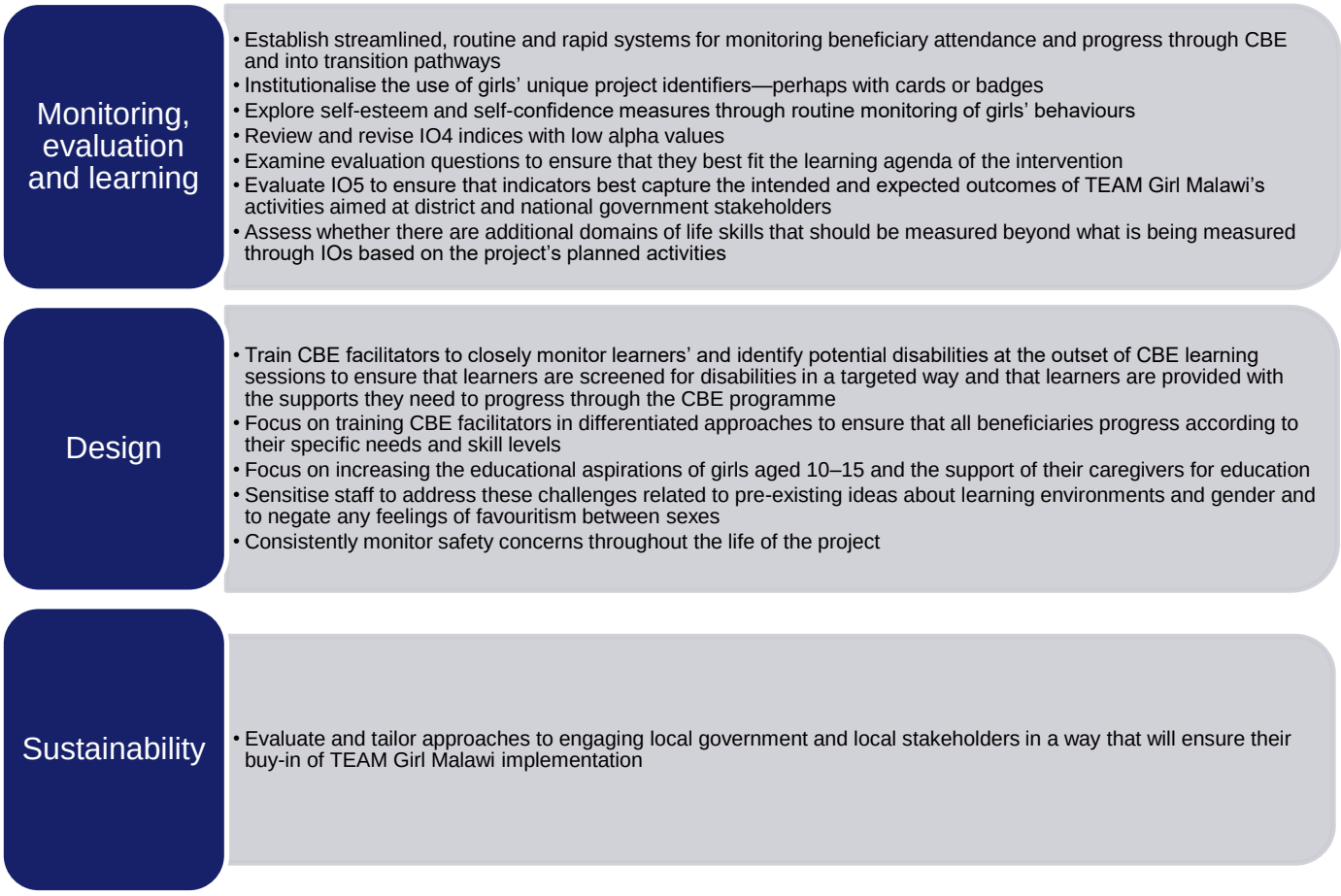
Summary baseline conclusions and the appropriateness of project interventions are described below.

- The TEAM Girl Malawi enrolment process conducted by the project effectively targeted girls within marginalised subgroups, except for girls who have albinism.
- Overall, TEAM Girl Malawi programming appears to be highly sensitive and proactive to responding to learning and transition barriers. SRHR, safety and community engagement in girls' education are all key themes of the project's ToC. Additional project inputs related to mitigate financial barriers—such savings groups' engagement and microloan disbursements— also target beneficiaries.
- Girls' baseline literacy levels are notably low, though girls' numeracy scores at baseline were more encouraging. It is unclear how the project will target learning sessions to the different literacy and numeracy levels or subgroups of learners, though given the distribution in performance, the project should consider taking a differentiated approach, particularly to mathematics teaching.
- Most of the girls in cohort 1 (84.55%) attended some school before enrolling in TEAM Girl Malawi. Nearly all girls in the baseline sample (94.40%) said they believed they would finish CBE, and more than one-half hoped to go to vocational training after finishing. A smaller proportion of girls—about one in 5—expressed a hope to return to primary school after completing CBE. Findings also indicate that girls of different ages have different expectations for transitions after CBE, which aligns with TEAM Girl Malawi transition pathways. Given that only about 1 of 3 girls aged 10-11 hope to return to primary school, the project should focus attention on changing perceptions of the opportunities gained through the formal school system.
- When evaluating the project's planned transition pathways through a GESI lens, the pathways appear to be gender accommodating. The project should ensure that it is encouraging inclusive education opportunities in the formal school system for girls with disabilities who choose to continue through primary school. Further, TEAM Girl Malawi should encourage vocational training and employment opportunities that are accessible to all girls, regardless of their functional difficulties. The project may consider how to mitigate the barriers faced by young mothers, who may be restricted from engaging in formal education and training opportunities due to their childcare responsibilities.
- Overall, evidence at baseline suggested mixed levels of enabling environments for sustainability. Further research into these indicators will be collected and analysed at the next evaluation point.
- TEAM Girl Malawi interventions fulfil the requirements of 'gender sensitive'. The project addresses many of the identified barriers that restrict girls' learning and transition, and it is aimed at changing perceptions throughout the communities where it is working. By including equally marginalised boys in CBEs, the project is aiming to reduce resentments and perceptions of favouritism while also allowing support for social-norm change and equality. The project should remain cognisant of gender dynamics in the classroom, especially as there will be a range of ages of girls and boys in the learning environment. Girls and boys should also be actively engaged together in Girls Club sessions, to ensure that sensitive topics can be explored in a gender transformative ecosystem.
- Assumptions in the ToC regarding subgroups and barriers appear to hold true. The most prevalent social, economic and educational barriers uncovered through the baseline are considered in TEAM Girl Malawi intervention planning. These include support for girls' SRHR—specifically menstrual health—through Girls' Clubs, financial support through

micro-loans for households with poverty or food insecurity and system-level support for families. TEAM Girl Malawi may want to revisit assumed educational barriers through monitoring to ensure that they continue to be applicable to the beneficiary population and communities.

Baseline recommendations are summarized in Figure 9.

Figure 9: Summary of baseline recommendations



3. Background to project

The Transformational Empowerment for Adolescent Marginalised Girls in Malawi (TEAM Girl Malawi) project is a 5-year Department for International Development (DFID)-funded Girls' Education Challenge (GEC) initiative through the Leave No Girl Behind (LNGB) funding window. Link Community Development (Link) implements TEAM Girl Malawi in collaboration with consortium partners Theatre for a Change (TfaC), Charlie Goldsmith Associates (CGA), Supreme and MicroLoan Foundation. School-to-School International (STS) serves as the external evaluator (EE) for TEAM Girl Malawi.

3.1 Project context, target beneficiary groups and theory of change

Project to complete

- Please outline:
    - The main contextual factors that have influenced the project design (e.g. political, economic, social, environmental, legal and/or educational policy/system context).
    - How gender inequalities and marginalisation impact the education of girls in these areas.
    - If the context is the same or different across all the areas the project is working (e.g. is one more rural? Does one area have higher poverty, different language or education system/policy? Etc.).
    - How your project defines its direct beneficiaries. This definition should include the main characteristics girls must have to be enrolled into your project. Please also ensure you discuss if any prioritisation criteria was used to select the most marginalised direct beneficiaries and if the project was oversubscribed.
    - If applicable, how the direct beneficiaries were selected for cohort one and how future cohorts will be selected.

- Complete Table 1, 2 and 3.

- Add your Project's latest ToC diagram in this document or as an annex and briefly summarise it, including the activities, intermediate outcomes, assumptions and barriers you're aiming to overcome.

Context for programme design

Politically, Malawi is stable. However, rising inflation, corruption, inequality and climate change leave 75% of people living under the poverty line (World Bank, 2015), and the country placed 170 on the Human Development Index 2016. The Ministry of Education, Science and Technology (MOEST) has inadequate funding and capacity, and the 2015-2016 Education Sector Performance Review indicates the country will not reach its education targets. Malawi continues to be dominated by traditional authorities' bylaws that often conflict with national laws, particularly around issues of child safeguarding.

In recent years, Malawi has experienced widespread drought and flooding leading to more than 50% of people experiencing food shortages. Poor rural people and children are particularly vulnerable to climate related shocks. Health care is weak—10.6% of the adult population are HIV

positive (Ministry of Health 2016), and the epidemic, combined with shortages of medical supplies, plays a strong role in the country's low life expectancy of 57 years for men and 60 years for women (WHO 2015).

Traditional sociocultural expectations place significant barriers on the ability of girls living in poverty to succeed educationally and economically. A 2016 UNICEF study found that 46% of girls marry, and 35% of girls give birth before age 18. Additionally, 20% of girls experience sexual violence, and exploitation and abuse remain accepted norms (Ministry of Gender, Children, Disability and Social Welfare, MOGCDSW 2015). Almost half, 47 percent, of girls complete primary education, compared with 56% of boys (EMIS 2015). The Child Protection (CP) system is under-resourced and weak.

In the Central Western Region of Malawi there are above average rates of girls' dropout, standards repetition, orphans and child headed households (EMIS 2015). Dedza's education system is overstretched due to the migration of children from Mozambique (NESP 2008–2017). Mchinji has a chronic lack of teachers, with almost no provision for children with special needs (NESP). In Lilongwe, there is particularly high risk of trafficking and sexual exploitation. The TEAM Girl Malawi project responds to the reality of this context.

A gender and social inclusion analysis informed the project design and theory of change (ToC). It also identified multiple and intersecting barriers that prevent highly marginalised girls from accessing a quality education.

Social marginalisation

- Early and forced marriage of girls is culturally accepted and provides income for poor families. It is rare for married girls to remain in school.
- Deeply ingrained attitudes denigrate girls' education as something of low value with little positive return. There remains a prioritisation of boys' education, heightened by the fact that girls are expected to take on more household chores and care responsibilities.
- Teenage pregnancy is common and increasing both for married and unmarried girls. Whilst the Readmission Policy is implemented in the target districts, girls report childcare, poverty, stigma and feeling 'too old' for school as reasons for dropping out. Young fathers are less likely to dropout.
- Gender-based violence and child abuse is normalised and common in both the school and community environment, and CP systems are weak. According to one study, 24% of children have experienced multiple forms of violence, with boys being more likely to experience physical while girls experience sexual violence (2013 VACS). Adolescent girls report feeling unsafe travelling to school.
- Malawi is a conservative country and adolescents who experience stigma from disability, HIV status, mental health, albinism or sexual exploitation are particularly vulnerable. This is compounded by poor access to health services and few schools providing an inclusive, safe environment. Girls remain at high risk of HIV—3.7% of young women aged 15–17 live with HIV compared to 0.4% of boys (MoH 2014).

Economic marginalisation

- Whilst primary school is free, families who suffer poverty are unable to afford essential additional costs of school— books, uniform, exam fees—and rely on income from child labour. This is particularly true for child headed households and orphans.

- Adolescent girls are at risk of sexual exploitation for income generation and from internal and external trafficking. It is challenging for a sexually exploited girl to return to school, particularly if she is contributing to the household income.
- In Lilongwe, there are additional challenges of slum living.

Educational marginalisation

- Primary schools are under-resourced, and teachers are unable to provide vulnerable children with individual attention and support. Gender norms mean that girls participate less than boys, which impacts their self-confidence as well as their ability to progress. Girls' learning is restricted by pedagogy that is not gender responsive. Primary schools are rarely equipped with separate sanitation facilities for girls and do not meet their needs during menstruation.
- MOEST (2009) acknowledges that teachers are ill-equipped to teach life skills.
- Adolescent girls are reluctant to re-join classes with younger children, where the pedagogy is inappropriate for their age.
- Despite a government policy to make available alternative forms of education for marginalised, vulnerable or over-age children, Malawi's provision of complementary-based education (CBE) is patchy, non-existent or mismanaged.
- Most, 59%, of the primary school teachers are male (EMIS 2015). Girls lack role models in the education sector, which becomes particularly challenging as they negotiate puberty and socio-cultural expectations.
- Low parental literacy levels, particularly of women, and few educational resources prevent children from accessing educational support at home.

Direct beneficiaries of the TEAM Girl Malawi Project are defined as those who are the intended, targeted beneficiaries of the interventions. The interventions are designed specifically to meet the needs of direct beneficiaries and support their vulnerabilities, and to tackle the barriers which they face to obtain basic levels of literacy/numeracy as well as being equipped to access SRHR rights, choice and safety.

Beneficiary selection for direct beneficiaries for cohort 1 took place using eligibility criteria that learners had to meet: be out of school, age 10–19 and have no literacy or numeracy skills. Vulnerability criteria are based on the identified barriers to education and were ranked in order of magnitude by the communities themselves before being used to ascertain the most vulnerable individuals. In general, across the 40 communities in the three districts, there was an over subscription by more than 50% (5006) of beneficiaries who were registered eligible for cohort 1. Therefore, the vulnerability criteria and its application process became very important. Communication was key not to create ill-feeling from those who would not make it into cohort 1. In the majority (approximately 75%) of the 40 communities, it is likely the project will hold cohort 2 in the same location, which will then be an option for those who could not attend cohort 1 to participate in the project. However, a fresh beneficiary selection process will take place for transparency. The targeting of cohort 3 will be determined through fresh and timely familiarisation and analysis of context and need in collaboration with district authorities, communities and other stakeholders.

Table 1: Summary of direct beneficiaries

Direct beneficiary numbers	Total figures
Total number of girls reached in cohort 1	2009
Total number of girls expected to reach by end of project	6000

Education level	Proportion of total direct beneficiaries (%)
Never been to school	286 (14.24%)
Been to school but dropped out.	1723 (85.76%)
Age banding (The age bandings used should be appropriate to the ToC)	**Proportion of total direct beneficiaries (%)**
10 to 15	747 (37.18%)
16 to 19	1262 (62.82%)

Table 2: Proposed intervention pathways after successful CBE completion

Intervention pathway	Which girls follow this pathway?	How many girls follow this pathway for cohort 1?	How long will the intervention last?	How many cohorts are there?	What literacy and numeracy levels are the girls starting at?	What does success look like for learning?	What does success look like for transition?
Enrol back into school (standard 5) (transition group A)	Girls aged 10–15 at end of 2 years of CBE	800	Ongoing	N/A	Standard 0–1 for literacy and numeracy	Girls achieve standard 4 equivalent for literacy and numeracy	Girls enrol back into school (standard 5) and continue learning
Embark on supported vocational training course (transition group B)	Girls aged 16–17 at end of 2 years of CBE	400	6 months	N/A	Standard 0–1 for literacy and numeracy	Girls achieve standard 4 equivalent for literacy and numeracy	Girls obtain skills to earn
Enter Micro Loan Group after business training and selection (transition group C)	Girls aged 18–19 at end of 2 years of CBE	400	Ongoing	N/A	Standard0–1 for literacy and numeracy	Girls achieve standard 4 equivalent for literacy and numeracy	Girls repay loan and continue with business earning

Table 3: Indirect beneficiary groups

Group	Interventions received	Total number reached for cohort 1
Boys	CBE curriculum	407
CBE facilitators and learning assistants, AoCs	Extensive training and job experience	80 facilitators 100 learning assistants 95 AoCs
Wider community members	Community sensitisation on numerous issues, such as inclusive education and tackling stigmatisation	10,600

Group	Interventions received	Total number reached for cohort 1
Family members of direct beneficiaries	Household economic benefit of vocational training, business training and loans	10560 (average household size of 4.4 x 2400)
District officials (PEAs, teachers, etc.)	Inclusion training in schools and capacity building	280

4. Baseline evaluation approach and methodology

The following section presents information on the baseline evaluation approach, including details on the overall evaluation purpose and questions, quantitative and qualitative methodologies, data collection tools, enumerator training and operational baseline data collection. The baseline was conducted by the TEAM Girl Malawi EE, STS, and the local data collection firm, the Centre for Educational Research and Training (CERT) at the University of Malawi.

4.1 Evaluation purpose(s) and evaluation questions

The overall purpose of the evaluation of TEAM Girl Malawi is to test assumptions that underpin the project's ToC. In other words, the evaluation is designed to provide relevant, meaningful and credible findings about the logical design of the project and its ability to meet its proposed outcomes in relationship to IOs.

TEAM Girl Malawi's primary and sub-evaluation questions and data sources are detailed in Table 4. Where baseline evidence is available, a report section is referenced. Four project-level evaluation questions guide all LNGB project, and these are further specified by the project-specific sub-evaluation questions. The sub-evaluation questions align with TEAM Girl Malawi's ToC and measure the implementation assumptions the project was designed on. Results for the sub-evaluation questions will be aggregated across the sample to answer the primary evaluation question.

Some evaluation questions will be answered using qualitative or quantitative evaluation-level data, while others will be answered using project-level monitoring data. STS and Link will work collaboratively to ensure that findings are presented in a fair and credible manner, even when data is collected directly by the project.

Table 4: Evaluation questions and summary of quantitative and qualitative data or analysis required to answer question

Evaluation question	Qualitative data or analysis required to answer question	Quantitative data or analysis required to answer question
1. What impact did the GEC funding have on the transition of highly marginalised girls into education, learning, training or work opportunities?		
1a. How do participating girls' learning and transition outcomes compare to those of nonparticipating girls? (Section 6.3)	Findings from focus group discussions (FGDs) with learners and caregivers and key informant interviews (KIIs) with CBE facilitators will be used to contextualise quantitative findings.	Early Grade Reading Assessment (EGRA), Early Grade Mathematics Assessment (EGMA) and transition data for cohorts 1 and 3 will be analysed to understand learning and transition outcomes between a

Evaluation question	Qualitative data or analysis required to answer question	Quantitative data or analysis required to answer question
		proxy comparison group (i.e. the benchmark group).[4]
1b. How do girls' mathematics and literacy performance vary by levels of IOs? (Section 6.2)	Findings from FGDs with learners, caregivers and CBE facilitators will be used to contextualise quantitative findings.	EGRA and EGMA data and IO data will be examined to assess performance by subgroups and to understand correlations between IOs and learning outcomes.
1c. How do the CBE exam pass rates for girls' and boys' compare? How do the transition rates for girls and boys compare for vocational training, microloan and primary school?[5]	Findings from FGDs with girl and boy learners will be used to contextualise quantitative findings.	CBE exit exam data and transition data will be examined by gender. Data on quality of education will be used to understand correlations.
2. What works to facilitate the transition of highly marginalised girls into education/training/employment and to increase learning?		
2a. To what extent were the TEAM Girl Malawi interventions adapted to address challenges faced? (Section 5.1)	Findings from KIIs with CBE facilitators will be used to contextualise quantitative findings.	Project monitoring data will be used to understand how transition rates were impacted by modifications or adaptations over the life of the project.
2b. To what extent has the GEC reached and affected highly marginalised girls? (Section 5.1)	Findings from FGDs with learners and caregivers and KIIs with CBE facilitators will be used to contextualise quantitative findings.	Project monitoring data will be used to understand how transition rates were impacted by learner marginalisation levels.
2c. Are appropriate levels of resources available to identify and support subgroups of extremely marginalised girls? [6]	Findings from KIIs with CBE facilitators, government officials and project staff will be used to understand resourcing levels and priorities over time.	NA
3. How sustainable were the activities funded by the GEC and was the programme successful in leveraging additional interest, investment and policy change?		
3a. To what extent are TEAM Girl Malawi activities embedded in CBE and MOEST and MOGCDSW processes, structure and staff capacities? (Section 6.4)	Findings from KIIs with CBE facilitators, government officials and project staff will be used to understand levels of embeddedness of activities.	Quantitative data from CBE facilitators will be used to understand processes and activities embedded in CBEs.
3b. Do communities demonstrate ownership over improving education for girls in TEAM Girl Malawi target areas? (Section 6.4)	Findings from FGDs with mothers' groups and caregivers and KIIs with community leaders will be used to understand community engagement in—and support and ownership of—girls' education initiatives.	Quantitative data from household and girls' surveys will be used to understand community engagement in—and support and ownership of—girls' education initiatives.

[4] Benchmark results will be available at the year 3 evaluation point. See Annex 6: MEL framework and Annex 7: External evaluator's inception report (where applicable) for more details.
[5] Findings will be available at the year 3 evaluation point.
[6] Findings will be available at the year 3 evaluation point.

Evaluation question	Qualitative data or analysis required to answer question	Quantitative data or analysis required to answer question
3c. To what extent has TEAM Girl Malawi leveraged additional resources (financial, human, in-kind) to support programme activities?[7]	Findings from KIIs with CBE facilitators, government officials and project staff will be used to understand levels of resources available for project activities.	Project monitoring data will be used to understand resource access and availability over time.
4. How successfully did LNGB projects reduce barriers to full participation in education or vocational education for highly marginalised girls?		
4a. How have TEAM Girl Malawi interventions affected girls' attendance, awareness of SRHR? (Section 7)	Findings from FGDs with learners and KIIs with CBE facilitators will be used to understand dosage of interventions and impact on attendance and SRHR awareness.	IO attendance data and data on SRHR will be used to understand change over time. Dosage data—defined as the amount of content on SRHR delivered—will be used for correlation analyses, if available.
4b. How have TEAM Girl Malawi interventions affected the quality of education at CBE Centres and Primary Schools? (Section 7)	Findings from FGDs with learners and caregivers and KIIs with CBE facilitators will be used to understand perceptions of and challenges to quality of education.	IO data will be used to understand change in education quality over time. Dosage data—amount of learning content—will be used for correlation analyses, if available.
4c. How have TEAM Girl Malawi interventions affected community attitudinal changes? (Section 7)	Findings from FGDs with mothers' groups and caregivers and KIIs with community leaders will be used to understand community attitudes on girls' education, gender and SRHR over time.	IO data will be used to understand changes in community attitudes on girls' education, gender and SRHR over time.
4d. How have TEAM Girl Malawi interventions strengthened leadership and engagement at school, district and national levels? (Section 7)	Findings from KIIs with government officials will be used to understand leadership and engagement on girls' education.	NA

4.2 Overall evaluation design

The evaluation of TEAM Girl Malawi project employs a mixed-methods, longitudinal, quasi-experimental design. The evaluation will utilise data from learning assessments and a package of quantitative and qualitative instruments from different respondents and, in subsequent timepoints, ongoing project monitoring tools. The variety of tools, respondents and methods of data collection will allow for the data to be triangulated and linked across evaluation questions and indicators.

Because TEAM Girl Malawi will roll out activities in a cohort design, and because of the ethical and logistical concerns in identifying a comparison group of girls for the evaluation, the evaluation will capitalise upon the cohort structures to benchmark findings against cohorts 1 and 3.[8] Evaluation data will be collected at 3 time points:

[7] Findings will be available at the year 3 evaluation point.
[8] As detailed in the MEL framework, TEAM Girl Malawi has determined that a comparison group is not appropriate in the programme context. No services would be offered to comparison group girls, which raises ethical concerns given levels of marginalisation. This could cause high levels of resistance from the community, MOEST and MOGCDSW. Further, these girls would be prohibitively difficult to track across evaluation points.

- Year 1 (July 2019): cohort 1 baseline
- Year 3 (July 2121): cohort 1 endline, cohort 3 baseline
- Year 5 (July 2023): cohort 1 follow-up, cohort 3 endline

A joint sampling approach will be used for the TEAM Girl Malawi evaluation. Specifically, STS and the project will collect learning and transition data for girls who are randomly sampled from cohorts 1 and 3. The team will also collect IO data from respondents—parents and caregivers, CBE facilitators, teachers, head teachers, community leaders—in the CBEs and communities where sampled girls live. Project monitoring data on attendance will be collected on a census-level by TEAM Girl Malawi and reported in subsequent evaluation reports.

The baseline evaluation design adheres to the current logframe and monitoring, evaluation and learning (MEL) framework. To examine the ToC's assumptions between IOs and outcomes, STS linked all data to girls' unique identifiers, allowing for analysis of the relationships between scores on IO indicators and outcomes. Additionally, the evaluation design is gender equality and social inclusion (GESI) transformative. The evaluation design considers gender, disability and other social differences and inequalities. These characteristics are explicitly accommodated in the selection of project beneficiaries, the design of evaluation tools and protocols for administration, the sampling of respondents, the selection and training of enumerators and the reporting of evaluation results. Although the project is inclusive of adolescent marginalised boys, quantitative baseline data was only collected from girls per the TEAM Girl Malawi MEL framework and STS' baseline inception report.

4.3 Evaluation ethics

STS adhered to TEAM Girl Malawi ethics, CP and safeguarding policies throughout the baseline process. This included providing all CERT staff and enumerators with relevant policies and engaging TEAM Girl Malawi to present on the policies during enumerator trainings. Enumerators were provided with TEAM Girl Malawi persons of contact for each district to ensure that any ethical issues that arose could be mitigated or reported. A summary of the ethical protocols and the baseline approaches to adhering to protocols is presented in Supplemental Table 1.

One ethical issue arose during the in-field practice during the quantitative enumerator training. An enumerator discovered an instance of child abuse and immediately reported the issue to the TEAM Girl Malawi staff on-site, which included the programme officer, MEL officer and CBE facilitators.

Supplemental Table 1: Ethical protocols and baseline approaches

Ethical issue/protocol	Baseline approach
Your overall MEL approach, including your evaluation design (including any use of control or comparison groups), your overall monitoring system and your approach to learning	STS adhered to TEAM Girl Malawi's MEL framework regarding the evaluation design. Specifically, no control group was included in the baseline due to the logistical and ethical issues identified by the project. TEAM Girl Malawi provided the sampling frame to STS, which included enrolled girls who were selected into the project based on their level of marginalisation. All marginalised groups were represented in the sample, apart from girls with albinism (see Section 5.1).
Quantitative and qualitative data collection methods and tools	Baseline tools included items related to SRHR, gender-based violence and child abuse that are sensitive in nature. STS ensured that all items were reviewed by Link and TfaC to ensure that items

Ethical issue/protocol	Baseline approach
	were inclusive and appropriate for the context and that Chichewa translations were responsive to the sensitiveness of the topics. All enumerators at baseline were female, and STS and TEAM Girl Malawi provided training and guidance to data collectors—both qualitative and quantitative—on administration of sensitive sections of the tools. Further, introductions to sections with sensitive items were added to prepare respondents for the types of questions they were asked. All respondents were given the option to refuse to respond to all items that included sensitive topics. Enumerators also received training on the selection of appropriate areas in which to administer surveys to ensure CP and ensure privacy.
Quantitative and qualitative sampling approaches	Qualitative sampling was conducted to ensure that all TEAM Girl Malawi subgroups were given the opportunity to participate and to capture perspectives and experiences of marginalised groups. Specifically, FGDs were conducted with homogeneous groups to encourage participants to voice their opinions in an inclusive and safe space. In addition to administering FGDs with all identified marginalised subgroups, FGDs were conducted with participants in age groups of 10–14 and 15–19, separately. This was done to create an environment in which younger girls and boys felt empowered to share their thoughts without the pressure of older participants' presence. Quantitative sampling was not stratified by subgroup or level of marginalisation. However, all marginalised subgroups were represented in the quantitative sample (see Table 13 for the proportion of girls from the sample in each characteristic subgroup).
Quantitative and qualitative data collection process, including your approach to seeking consent or assent	All enumerators received TEAM Girl Malawi ethics, CP and safeguarding policies. They also received training directly from TEAM Girl Malawi on the policies and on protocols for reporting violations of policies or instances of breaches of CP. Enumerators read a consent or assent statement to respondents prior to initiating the learning assessments, girls surveys and household surveys. These statements included all information commonly required by institutional review boards and allowed respondents to voluntarily end their participation, without penalty, at any time. Further, at the beginning of sections with sensitive items on the girls and household surveys, respondents were read a statement about the types of questions that would be asked and were reminded that they could choose not to answer any questions without penalty.
Recruitment, training and supervision of MEL personnel	STS and CERT recruited an all-female team of enumerators who underwent background police checks in line with TEAM Girl Malawi's recruitment policy. They also underwent all required trainings on inclusion, ethics, sensitivity and CP per TEAM Girl Malawi policies. All data collectors signed confidentiality and data protection statements at the beginning of training.

Ethical issue/protocol	Baseline approach
	During data collection, STS maintained contact with supervisors through WhatsApp to ensure that no ethical issues arose. All supervisors also had contact information for TEAM Girl Malawi staff.
Data recording, storage, analysis and reporting	All baseline data was collected using password protected software programmes. Qualitative notes were returned by enumerators to CERT at the end of data collection for safe storage, and all software programmes were deleted from tablets at the end of quantitative data collection. STS securely downloaded and stored all baseline data on password protected servers. Public-use files were created to share baseline data, and all data presented in the baseline report are anonymised.

4.4 Quantitative evaluation methodology

Quantitative evaluation tools

Two baseline evaluation surveys and 2 learning assessments were developed and used for the quantitative component of the evaluation. The development of the learning assessments for TEAM Girl Malawi are described in additional detail in the corresponding sections. STS and TEAM Girl Malawi collaboratively developed the survey tools, detailed in Table 5, prior to pretesting and data collection. They include a girls survey and a household survey. The tools combined numerous domains relevant to the project's ToC and items that corresponded to the project's logframe indicators. Each tool uses LNGB templates as the initial source of items. Following the compilation of these items and additional project-specific items within each tool, STS shared drafts with TEAM Girl Malawi and TfaC, who commented and provided revised or new items based on the project's indicators and specific implementation priorities. All item sources and revisions were tracked in a master file. Both surveys were shared with the fund manager (FM) for review and approval prior to the pretest and operational data collection.

Table 5: Quantitative evaluation tools (baseline)

Tool name	Relevant indicator(s)	Who developed the tool?	Was tool piloted?	How were piloting findings acted upon (if applicable)	Was tool shared with the FM?	Was FM feedback provided?
Girls survey	O1.3 O2 IO2.1 IO2.2 IO3.1 IO4.1 IO4.2 IO4.3	STS, Link, TfaC	Yes – pretested	Minor modifications to translations and problematic items made following pretest	Yes	Yes
Household survey	IO4.2 IO4.3	STS, Link, TfaC	Yes – pretested	Minor modifications to translations and	Yes	Yes

Tool name	Relevant indicator(s)	Who developed the tool?	Was tool piloted?	How were piloting findings acted upon (if applicable)	Was tool shared with the FM?	Was FM feedback provided?
				problematic items made following pretest		
EGRA	IO1.1	STS (adapted from existing tools)[9, 10]	Yes – pretested	Significant updates made to reading passage and listening passage to align with quality guidance and to make gender appropriate	Yes	Yes
EGMA	IO1.2	STS (adapted from existing tools)[11]	Yes – pretested	Minor modifications to translations and examples added following pretest	Yes	Yes

It is expected that the 2 surveys should remain relatively stable across the evaluation points, with minor revisions or additions required.[12] However, different and equated forms of the learning assessments will be administered at midline and endline.[13] Additional quantitative tools will be developed for midline and endline evaluation points to measure indicators that did not require baseline values. These will likely include CBE facilitator surveys, head teacher surveys and classroom observation forms.

Enumerators

STS and CERT worked collaboratively to recruit, hire and train enumerators for the pretest and operational baseline data collection activities. STS provided CERT with a list of key qualifications and job descriptions, and CERT recruited local female enumerators who fit the required qualifications. Following initial screenings, oral interviews and reference checks, CERT selected 15 enumerators for the quantitative activity and distributed their curriculum vitaes to STS for final review. All selected enumerators had prior experience conducting surveys, either on paper or electronically, and nearly half had experience conducting EGRAs using Tangerine®, an open-source software developed by RTI International. All were fluent in Chichewa.

[9] Creative Associates International, RTI International and Seward Inc. *Malawi National Early Grade Reading Assessment Survey: Final Assessment – November 2012*. Washington, DC: USAID, 2012.
[10] USAID/Malawi and MOEST. *USAID Funded Malawi Teacher Professional Development Support (MTPDS) Activity 2010 Early Grade Reading Assessment (EGMA): National Baseline Report 2010.* Washington, DC: USAID, 2010.
[11] USAID/Malawi and MOEST. *USAID Funded Malawi Teacher Professional Development Support (MTPDS) Activity 2010 Early Grade Mathematics Assessment (EGMA): National Baseline Report 2010.* Washington, DC: USAID, 2010.
[12] This assumes that the programme's ToC also remains stable across evaluation points. Revisions or additions will be based on learnings from the baseline and implementation.
[13] Equating data was captured at baseline but will be reported in the year 3 evaluation point report.

Before training commenced, all selected enumerators signed contracts with CERT that stipulated their expected roles, including their expected ethical and professional conduct during training and data collection. Additionally, all enumerators underwent police security clearance checks as required by Link as part of its child safety and protection procedures for all persons working under their projects.

Three of the 15 enumerators were selected to participate in the quantitative pretest, which took place from 28–30 May 2019 in Lilongwe. On day 1 of the pretest, STS trained the enumerators on the 4 quantitative evaluation tools. On day 2, enumerators administered learning assessments to 20 respondents, girls' surveys to 8 respondents, and household surveys to 5 respondents. On day 3, enumerators provided feedback on their experience and on specific components of the tools. Their feedback was incorporated into the revisions presented to Link and the FM prior to the start of operational data collection.

The baseline quantitative enumerator training, facilitated by STS with support from CERT and Link, took place from 1–5 July 2019 in Lilongwe. During the training, enumerators were split into 2 groups—those responsible for administering surveys and those responsible for administering the learning assessments. STS based group assignments on the enumerators' previous experience and expertise. Sessions were delivered in plenary and group formats and included the following topics:

- Baseline study purpose and research ethics
- Introduction to TEAM Girl Malawi project
- Safeguarding and CP
- EGRA/EGMA and equating tests
- Surveys
- Using tablets for data collection
- CBE mobilisation and team roles and responsibilities
- Accommodations for girls with disabilities
- Data collection logistics
- Supervisor roles and responsibilities

Learning assessment enumerators took part in 2 assessor accuracy quizzes during the training, through which it was possible to measure enumerators' ability to score consistently and accurately with a 'gold standard', or a script of responses. All enumerators scored over 90% on both quizzes, indicating high assessor accuracy. The training schedule also included one day of in-field practice, during which 2 groups of enumerators visited 2 different TEAM Girl Malawi CBE communities that were not part of the baseline sample.

On the last day of training, CERT divided the enumerators into 3 teams, each consisting of 3 enumerators and one supervisor-enumerator. All teams had 2 enumerators trained in surveys and 2 trained in learning assessments. On the same day, STS led a supervisor training to ensure that the 3 supervisors, 2 of whom are CERT staff, were aware of their roles and responsibilities during and after data collection.

Quantitative data collection

Quantitative data collection took place from 8–18 July 2019. Each CBE visit spanned 2 days. Team A was assigned to CBEs in Dedza, Team B to Dedza and Lilongwe and Team C to Mchinji.

During the visit, teams collected a quota of 27 learning assessments, 27 girls surveys and 27 household surveys.

All data was collected electronically on Android-based tablets. The learning assessments were administered to girls using Tangerine®, and surveys were administered using SurveyCTO, a mobile data collection platform. At the end of each day, supervisors upload all data from their team's tablets to the software servers, and STS' quality control team downloaded and securely stored all raw data on a password-protected server for review, cleaning and analysis. After data collection was completed, CERT ensured that the software and any TEAM Girl Malawi data was permanently deleted from the tablets and that all paper documents with identifying information were discarded.

STS assured data quality through several strategies. The use of tablets for electronic data capture mitigated data entry errors and helped ensure data quality, consistency and collection efficiency. Records were linked across tools using TEAM Girl Malawi's beneficiary unique beneficiary identifiers, which were programmed into all tools and populated into the dataset. During CBE visits, supervisors completed tracking sheets to keep record of girls who had been assessed, girls who completed the girls survey and parents or caregivers who completed the household survey. At the end of the CBE visit, supervisors used the tracking sheets to complete a control form in SurveyCTO, which was cross referenced by STS' quality control team with the sample and TEAM Girl Malawi's enrolment database. As a result, it was possible to confirm daily which and how many tools were completed, determine any data quality issues and ensure that the correct girls were sampled. Any issues or challenges were recorded into a data collection tracker. STS' quality control team coordinated directly with supervisors through WhatsApp to reconcile any quality issues.

Quantitative data cleaning and storage

STS stored all raw data on a password-protected server. Raw datasets underwent 3 levels of data cleaning based on a standard protocol. During level 1, final raw data was reviewed and flagged for duplicates, inappropriate time and date submissions, inacceptable administration lengths, inconsistency in CBE samples compared to the expected, revoked and refused consent and missing data. At level 2, specific disposition codes, taken from the quality control team's data collection tracker, were integrated to the dataset to remove or adjust cases based on issues uncovered during the data collection.[14] After level 2, STS again reviewed datasets for duplicates, missing data and inconsistencies to ensure all issues were resolved. Finally, at level 3, STS computed learning assessment subtask scores, aggregate literacy and numeracy scores and survey composite scores. Outliers were identified and examined for inconsistencies. At the end of the 3 levels of cleaning, datasets were merged to complete the analysis.

Quantitative data analysis

All quantitative data were analysed using Stata and IBM SPSS® software platforms. The learning assessment analysis included girls who were sampled and who had unique identification numbers that matched the TEAM Girl Malawi enrolment database. The raw learning assessment data included 387 records with data and affirmative consent. Of these, 7 cases were dropped because they were duplicates, resulting in 380 remaining records.[15] From the remaining 380, 11 duplicate

[14] Disposition codes are STS' internal system for data cleaning. Specifically, disposition codes indicate the type of issue in a record or data point as well as the proposed resolution. During the cleaning process, disposition codes assist the analyst to determine the extent of discrepancies in a specific record or variable and make appropriate decisions about the data quality and cleaning.
[15] Duplicates occurred when one assessment was initiated and incomplete for a girl, and a subsequent record was completed for the same girl.

unique identification numbers remained, 10 of which were cross-checked with TEAM Girl Malawi staff and corrected for retention. Two records were identified as girls who were assessed but were not in the project; those records were dropped. The final analytical learning assessment file contains 378 girls' data.

Similarly, the girls survey analysis included girls who were sampled and had a unique identification number that matched the enrolment database. The raw girls survey data included 373 records with data from a valid timeframe of data collection and affirmative consent. Records were dropped in the case of duplicate interviews and girls who were interviewed and not part of the project. The final girls survey analysis file contains 361 records.

Household survey analysis includes parents or caregivers of girls who were sampled and had a unique identification number that matched the enrolment database. The raw household survey data file contained 359 records from sample and replacement girls' households. One record was dropped because the household member was interviewed twice, and another was dropped because the respondent did not have a daughter who was part of the project. Extra records were created for caregivers who were surveyed once but had multiple beneficiaries in the project. In these cases, the caregiver's record was duplicated and tagged with the unique identification number of the second girl in the project. Ultimately, there were 360 records for unique girls in the household survey dataset and a total of 353 household survey respondents.

The girls and household datasets and the TEAM Girl Malawi enrolment database were merged to enable analysis of marginalisation characteristics and barriers to education (see Section 5). Finally, these datasets were merged with the learning assessment dataset. Out of the 378 learning assessment records, 11 were missing a girls survey and 21 were missing a household survey. Overall, 349 had both a girls and household survey.

All results use the unit of analysis that most accurately reflects the way in which the data were collected, and the items were structured. For all learning data, results are presented across girls, as the unit of analysis is the individual learner. For survey data, the unit of analysis varies. For indices related to aspects of a household, the unit of analysis is the respondent but is described as the household. For indices related to aspects of the community, the unit of analysis is respondents but is described as the community.

For the learning assessment, scores and learning bands were computed and reported per LNGB guidance. Guidance for aggregate scoring at year 3 may be revised to account for fluency rates on timed subtasks, instead of reporting only percentage correct.[16]

Composites—or indices—for IO indicators were created by mapping survey items to indicators. Relevant but non-overlapping items from the girls and household surveys were included in indices constructed for each indicator (Supplemental Table 2; see also Section 7).[17] Although the majority of indices were constructed based on the theory underlying the survey construction, the reliability of each composite was also checked by computing Cronbach's alpha.[18] For at least 2 of the composites, the alpha value was very low, indicating that either there were too few items comprising the composite, the items were not clearly understood by respondents or the items in the composite did not adequately measure the underlying construct. Additionally, the low reliability at baseline may also be due to incompleteness of data on all underlying items. While these composites were primarily constructed based on the design of the survey items—and alternative

[16] The FM will provide additional guidance on scoring at midline based on conversations with the funder.
[17] Only respondents who answered 25% or more of the underlying items were included in the index calculation.
[18] Chronbach's alpha is a measure of internal consistency or scale reliability. It measures how closely related a set of items are within a defined group.

items were not always available—revised items, additional items or additional response options may be necessary to improve reliability at midline.

Supplemental Table 2: IO Indices

Indicator	Index	Calculation	Number of Items	Scale range	Reliability
IO2.1: Number of girls with improved understanding of SRHR	SRHR understanding	Sum	18	0–18	0.69
IO2.2: Number of girls with improved self-esteem, self-confidence and well-being[19]	Self-esteem	Average	10	0–3	0.69
	Self-confidence	Average	4	0–3	0.69
IO4.2: Improved community support for SRHR and CP	SRHR support	Average	6	0–4	0.35
	CP support	Average	4	0–3	0.59
IO4.3: Improved community support for girls' education through CBEs and primary school	Girls' education support	Sum	12	0–15	0.43
O1.3: Number of highly marginalised girls supported by GEC with improved life skills outcomes	Life skills	Average	27	0–3	0.79

STS used regression models to understand the relationship between girls' scores on IOs and their marginalisation characteristic and barrier subgroups and demographics. The model examines the relationship between the independent variables—marginalisation characteristics, barrier subgroups and demographics—and the dependent variables—scores on IOs. The model also examines the relative importance of these factors in predicting scores on the IOs.

To determine the variables to include in the models while maintaining statistical power, STS explored girls' demographics and bivariate correlations between marginalisation characteristics and barriers. Demographic variables of age and district were included because age and district overlap—or are confounded—with the marginalisation characteristic and barrier subgroups.[20, 21] Additionally, the variable for a food insecurity or hunger barrier (see Section 5.2) was included as a proxy for low socioeconomic status since all girls in the sample fell into the high poverty subgroup. The remaining barriers and marginalisation characteristics had little or no correlation with the IOs. They were excluded from the regression models.[22]

[19] In conversations with TEAM Girl Malawi prior to the baseline, the team indicated that well-being would be removed from Indicator IO2.2. As a result, data for this domain was not captured at the baseline evaluation point.
[20] Specifically, results showed that age and district had statistically significant correlations with marriage, caregiving, chore hours and starting menstruation.
[21] Moderate correlations were observed with age. Weaker correlations were observed with district. See correlation results in Supplemental Table 44 and Supplemental Table 45. If correlated variables are included into a regression model together, the resulting findings may be unreliable.
[22] Some barriers and characteristics were excluded from the core set of predictors in the regression models because of low correlations with the IOs. Two predictors that had some relationship with some of the IOs were functional difficulty—whether a girl had 'a lot of difficulty' or 'could not do at all' when asked the Child Functioning questions on disability—and bullying—whether the

As a result, the analytical model includes age groups, district and food insecurity or hunger barriers.[23, 24] By including these demographic variables, the model examines the relative importance of the subgroups of interest—marginalisation and barriers—above and beyond their overlap with demographic variables.

Learning assessments

STS adapted learning assessments from existing EGRAs and EGMAs that had been previously administered in Malawi under the United States for International Development (USAID) Malawi Teacher Professional Development Support Programme, in collaboration with MOEST.[25] Both the EGRA and EGMA were administered in Chichewa, and the EGRA tested reading skills in Chichewa. Chichewa was selected as the assessment language because it is the national language of Malawi and the primary language of instruction through standard 4.

Details of EGRA and EGMA subtasks are included in Supplemental Table 3. Most subtasks included autostops—or early stop rules. This allowed enumerators to automatically stop one subtask and move on to the next if learners were unable to correctly answer a predetermined set of items. Autostops were established to allow learners to efficiently move through the assessment and to not spend a lengthy period trying to demonstrate skills that they do not have. Autostops also allowed for respondents with low learning levels to be exempt from attempting all items on each subtask. The length of time allocated for each timed subtask is noted in Supplemental Table 3.

Supplemental Table 3: Learning assessments

Tool name	Subtask	Purpose	Administration	Scoring
EGRA	Initial sound identification	Phonemic awareness	Untimed; autopstop after first 5 items	Correct initial sounds out of 10
	Letter name identification	Alphabet knowledge	Timed – 2 minutes; autostop after first 10 items	Correct letter names per minute; 100 items total
	Syllable identification	Alphabet knowledge and decoding	Timed – 2 minutes; autostop after first 10 items	Correct syllable sounds per minute; 100 items total
	Familiar word reading	Sight-word recognition and decoding	Timed – 2 minutes; autostop after first 5 items	Correct familiar words per minute; 50 items total

girl was bullied by peers or teachers. Functional difficulty was a predictor of girls' self-esteem and life skills. Bullying was a predictor of self-confidence and life skills. Additionally, girls who had high chore burdens were comparable to girls who had low chore hours. As a result, this variable was excluded from further analysis to maintain statistical power.

[23] Girls who were older were more likely to report having started menstruation. As a result, these two variables could not be included in a regression model together to avoid multicollinearity. Instead, in a particular regression model, the effects of age group on the outcome were examined. If no statistically significant relationship was found, menstruation was included in the model instead. Generally, results did not differ for the models when these two variables were used interchangeably, except for the self-confidence and support for girls' education indices. In these cases, menstruation was included rather than age.

[24] ANOVA showed that within district, statistical differences occurred between Mchinji and the two other districts but not between Dedza and Lilongwe. Therefore, including Dedza in the regression model reduced power and thus was excluded from the final model.

[25] The Malawi Teacher Professional Development Support activity was implemented by Creative Associates International, RTI International, and Seward Inc. from 2010 to 2013.

Tool name	Subtask	Purpose	Administration	Scoring
	Oral reading fluency	Decoding and reading fluency	Timed – 2 minutes; autostop after first 6 items	Correct words per minute; 54 items total
	Reading comprehension	Reading comprehension	Untimed; number of questions asked corresponds to how many words read in oral reading fluency passage	Correct out of 5
	Listening comprehension	Oral language comprehension and vocabulary	Untimed; all questions asked of all respondents	Correct out of 5
EGMA	Number recognition	Numerals and numericities identification	Timed – 2 minutes; no autostop	Correct per minute; 20 items total
	Quantity discrimination	Numerical magnitudes comparisons	Untimed; autostop after 4 consecutive incorrect items	Correct out of 10
	Missing numbers	Number patterns identification	Untimed; autostop after 4 consecutive incorrect items	Correct out of 10
	Addition (level 1)	Arithmetic skills	Timed – 2 minutes; no autostop[26]	Correct per minute; 20 items total
	Addition (level 2)	Arithmetic skills	Untimed; no autostop; only administered if respondent correctly answered at least one item correct on addition level 1 subtask	Correct out of 5
	Subtraction (level 1)	Arithmetic skills	Timed – 2 minutes; no autostop	Correct per minute; 20 items total
	Subtraction (level 2)	Arithmetic skills	Untimed; no autostop; only administered if respondent correctly answered at least one item correct on subtraction level 1 subtask	Correct out of 5

[26] Additionally, learners who did not correctly answer any items on the addition or subtraction level 1 subtasks were not asked items from the corresponding level 2 subtask.

Tool name	Subtask	Purpose	Administration	Scoring
	Word problems	Conceptual and real-word mathematics understanding	Untimed; autostop after 4 consecutive incorrect items	Correct out of 6

To best accommodate the TEAM Girl Malawi target beneficiaries—including those with disabilities, who are estimated at one-third of the sample (see Table 10), and those who may not have had recent classroom or test-taking experience, estimated at about one out of 7 (see Table 20)—all respondents were given 120 seconds on timed subtasks. This is a deviation from general EGRA and EGMA guidance, which stipulates that subtasks are timed at 60 seconds. STS and TEAM Girl Malawi determined that 2 minutes was an appropriate way to provide an inclusive assessment environment and to allow all girls an opportunity to acclimate to the testing environment and demonstrate their skill levels, regardless of their prior experience.

Additionally, enumerators had a set of large-print stimuli to provide to learners with low vision. STS also outlined set of recommended assessment accommodations for girls with disabilities that enumerators could utilise at their discretion. These included:

- Allowing a companion or supporter to accompany the girl during assessment for speech interpretation
- Allowing the girl to take breaks in-between subtasks
- Arranging data collection set-up to accommodate the girl if hearing is impaired on a specific side
- Pointing to items for the girl during assessment if she has a physical disability or mobility challenges

Enumerators made note of accommodations, such as assistive devices or large-print stimuli, that girls used at baseline in the data collection software; identical accommodations for each respondent will be provided at the next evaluation point to allow for consistent assessment conditions.[27]

Quantitative sample selection

Link constructed and populated the sampling frame for TEAM Girl Malawi using data collected during the beneficiary recruitment and enrolment process. The sampling frame included a unique identifier for each enrolled girl, as well as her location and demographic details. Further, the sampling frame included data on beneficiaries' marginalisation criteria that were established by the project (see Section 5.1), with the exception of albinism.[28] These criteria correspond to the project's characteristic subgroups of interest.

STS conducted sample selection for the cohort 1, joint-evaluation sample using a 2-stage stratified random sampling procedure. Guided by power calculations provided in the project MEL framework and STS' baseline inception report, the cohort 1 sample included a first-stage random selection of 14 CBEs, proportional to the total number of CBEs in cohort 1 (Supplemental Table 4). Specifically, 3 CBEs were randomly selected from Lilongwe, 7 from Dedza and 4 from Mchinji.

[27] In the final sample, 9 girls were provided with large print font. Zero girls were provided assistive devices such as glasses, magnifiers or hearing aids.

[28] TEAM Girl Malawi enrolment data did not capture information on albinism. The programme indicated that only one enrolled girl has albinism, and the programme excluded this information from their intake form due to the highly sensitive nature of the condition.

Following selection of the CBEs per district, STS randomly selected 27 girls and 5 replacements from each selected CBE, per the second stage of the sampling procedure. Based on the MEL framework and inception report, the only stratification variable accounted for during selection was age group—a decision made to ensure large enough sample sizes for cohort 1 and 3 benchmark comparisons.[29] Age stratification for the intended proportions of girls were:[30]

- 25% aged 10–11
- 50% aged 12–16
- 25% aged 17–19

The same girls in the learning assessment sample also form the girls survey sample. One parent or caregiver per sampled girl comprises the sample for the household survey.

CBE facilitators were responsible for mobilising the 27 girls and their caregivers to the assessment site for data collection.[31] If the selected girls were unavailable, the CBE facilitator contacted the 5 randomly selected replacements and their caregivers. If the quotas were still unmet, CBE facilitators recruited any other available beneficiary girls to the assessment site.

Quantitative sample sizes

Table 6 outlines the planned and actual sample sizes for the baseline evaluation.[32] Supplemental Table 4 details the sample sizes at the district level, using a probability proportional to size sampling procedure. At baseline, the learning assessment sample size was equal to the planned sample size, while the sample size for the girls survey and household survey were slightly less—response rates were 95.50% and 93.39%, respectively. Lower response rates were most often due to the availability of the respondents. When girls' availability was limited, enumerators prioritised administering the learning assessments over the girls survey.

Table 6: Quantitative sample sizes

Tool name	Sample size agreed in MEL framework	Actual sample size[33]	Remarks on why anticipated and actual sample sizes are different
EGRA/EGMA learning assessments	378	378	NA
Girls survey	378	361	95.50% response rate. Girls who were unable to dedicate additional time to the data collection or left the data collection site without being surveyed

[29] The quantitative sample age groups, as outlined in the MEL framework and inception report, are not the same as the transition pathway age groups. Age groups based on the quantitative sample are presented throughout the report, while age groups based on transition pathways are presented solely in Section 6.3.
[30] Due to low proportions, all girls in age group 10 through 11 were selected from sampled CBEs in Dedza and Mchinji. Girls were randomly selected in all other age groups in those districts. Girls were randomly selected into each age group for Lilongwe CBEs.
[31] CBE facilitators notified sampled girls and households of the dates and requirements for the data collection, encouraged and motivated them to attend and, in some cases, ensured their transport to and from the data collection site.
[32] The power calculations used to calculate the quantitative sample size are presented in Annex 6: MEL framework. Power calculations were computed to have statistical power at the aggregate level, not at the subgroup level. Statistical comparisons at the subgroup level presented throughout this report have lower levels of confidence.
[33] Actual sample size is representative of the number of records after data cleaning.

Tool name	Sample size agreed in MEL framework	Actual sample size[33]	Remarks on why anticipated and actual sample sizes are different
Household survey	378	353	93.39% response rate. Caregivers who were unavailable, absent, deceased or had multiple girls in the project[34]

Supplemental Table 4: CBE sample sizes

District	Total cohort 1 CBEs (% of total)	Total cohort 1 girls (% of total)	Sampled cohort 1 CBEs (% of total)	Sampled cohort 1 girls (% of total)
Lilongwe	8 CBEs (20.00%)	401 girls (19.97%)	3 (21.43%)	77 (20.37%)
Dedza	19 CBEs (47.50%)	951 girls (47.36%)	7 (50.00%)	193 (51.06%)
Mchinji	13 CBEs (32.50%)	656 girls (32.67%)	4 (28.57%)	108 (28.57%)
Total	40 CBEs (100%)	2008 (100%)	14 (100%)	378 (100%)

Representativeness of the sample

Demographics of the baseline sample are presented in Tables 7 through 10. The representativeness of the baseline sample has been assessed by comparing Tables 7 through 10 with Supplemental Table 4 and the tables in Annex 5: Beneficiaries table (Project mapping data). Overall, the baseline sample is satisfactorily representative of the total beneficiary population.

Supplemental Table 4 and Table 8 provide details on the baseline sample and population breakdown by district. The baseline sample represents the TEAM Girl Malawi beneficiary population by district and age group, with results generalisable to the project level. In other words, the sample mirrors the population proportions by district and by age group.

By district, Dedza represents one-half of the sample and the population of TEAM Girl Malawi beneficiaries (sample: 51.06%, population: 47.36%). Mchinji represents one-third (sample: 28.57%, population: 32.67%) and Lilongwe represents one-fifth (sample: 20.37%, population: 19.97%).

Table 9 and Table 29 provide breakdowns of the baseline sample and beneficiary population by age. Girls aged 10-11 were intentionally oversampled at baseline. The sample was comprised of 15.88% of girls aged 10-11, while the population was comprised of 10%.[35] The proportions of girls in the sample and population were similar for those aged 12-16 (57.14% and 60%, respectively) and for those aged 17-19 (26.98% and 24%, respectively).

[34] Where one caregiver had multiple girls in the programme, their responses were imputed for all girls. This was the case for three caregiver responses.
[35] Girls aged 10 through 11 were oversampled because they will be compared to girls in cohort 3 at subsequent timepoints and will be of comparable age. See Annex 6: MEL framework.

Table 7: Sample breakdown by intervention pathways[36]

Intervention pathway	Sample proportion of intervention group (%)
Transition group A (girls aged 10–15 at end of 2 years of CBE)	28.04%
Transition group B (girls aged 16–17 at end of 2 years of CBE)	29.89%
Transition group C (girls aged 18–19 at end of 2 years of CBE)	42.06%
Source: N = 378	TEAM Girl Malawi enrolment database

Table 8: Sample breakdown by regions

Region	Sample proportion of intervention group (%)
Lilongwe	20.37%
Dedza	51.06%
Mchinji	28.57%
Source: N = 378	TEAM Girl Malawi enrolment database

Table 9: Sample breakdown by age

Age	Sample proportion of intervention group (%)
Aged <10 (%)	0.00%
Aged 10 (%)	7.94%
Aged 11 (%)	7.94%
Aged 12 (%)	2.65%
Aged 13 (%)	9.52%
Aged 14 (%)	12.43%
Aged 15 (%)	17.46%
Aged 16 (%)	15.08%
Aged 17 (%)	10.58%
Aged 18 (%)	16.40%
Aged 19 (%)	0.00%
Aged 20 + (%)	0.00%
Unknown	0.00%
Programme-specific age groups[37]	
Aged 10-11	15.88%
Aged 12-16	57.14%
Aged 17-19	26.98%
Source: N = 378	TEAM Girl Malawi enrolment database

It is not possible to fully assess the representativeness of the sample on disability prevalence. Beneficiary enrolment disability information was collected using the Washington Group Short Set of Disability Questions, while baseline disability prevalence was collected using the Washington Group/UNICEF Module on Child Functioning. Table 10 indicates that the proportion of the girls from the baseline with at least one domain of functional difficulty was 34.39%,[38] while the

[36] Transition pathway group A includes girls aged 10-13 at baseline, transition pathway group B includes girls aged 14-15 at baseline and transition pathway group C includes girls older than 15 at baseline.
[37] Age groups identified in the MEL framework and inception report and represented in the quantitative sample selection.
[38] If a girl responded, 'cannot do at all' or 'a lot of difficulty' on the Child Functioning questions, she was categorised as having a functional difficulty in each domain. See Supplemental Table 33.

proportion of enrolled girls with at least one domain of functional difficulty was 8.47%.[39] Given that the question sets and methodologies differ between the 2 sources, analysts cannot compare the sample proportions to the baseline population. Results on the Child Functioning questions are used for all baseline reporting.

Table 10: Sample breakdown by disability

Domain of difficulty	Sample proportion of intervention group (%)	Guidance – record as true if they meet the criteria below
Seeing	3.17%	If CF1=1 AND (CF2=3 OR CF2=4) **OR** If CF1=2 AND (CF3=3 OR CF3=4)
Hearing	3.17%	If CF4=1 AND (CF5=3 OR CF5=4) **OR** If CF4=2 AND (CF6=3 OR CF6=4)
Walking	5.82%	If CF7=1 AND (CF8=3 OR CF8=4) OR (CF9=3 OR CF9=4) **OR** If CF7=2 AND (CF12=3 OR CF12=4) OR (CF13=3 OR CF13=4)
Self-care	1.59%	CF14=3 OR CF14=4
Communication	2.38%	CF15=3 OR CF15=4 **OR** CF16=3 OR CF16=4
Learning	6.08%	CF17=3 OR CF17=4
Remembering	9.79%	CF18=3 OR CF18=4
Concentrating	6.61%	CF19=3 OR CF19=4
Accepting Change	6.08%	CF20=3 OR CF20=4
Controlling Behaviour	7.41%	CF21=3 OR CF21=4
Making Friends	5.82%	CF22=3 OR CF22=4
Anxiety	12.70%	CF23=1
Depression	11.38%	CF24=1
Girls with disabilities overall	34.39%	
Source: N = 378	Girls survey, household survey	

Challenges in baseline data collection and limitations of the evaluation design

STS and TEAM Girl Malawi faced several key challenges during the quantitative data collection and analysis:

- Some girls and caregivers initially selected into the sample were unavailable during data collection. The highest replacement rate was in Dedza among 12 through 16-year-old girls (17 replacement girls) followed by the same age group in Mchinji (10 replacement girls) and in Dedza among 17 through 19-year-old girls (9 replacement girls). Overall, 47 replacement girls took the learning assessment. The highest replacement rate for the girls survey was among 12 to 16-year-old girls in Dedza (26 replacement girls) and in the same

[39] If a girl responded, 'cannot do at all' or 'a lot of difficulty' on the Short Set questions, she was categorised as having a functional difficulty in each domain. See Supplemental Table 33.

age group in Mchinji (21 replacement girls). Overall, 85 replacement girls responded to the girls survey. Most girls recruited as replacements are project beneficiaries and therefore are retained in the baseline sample.

- Replacement girls were not matched 1-to-1 to the girl in the original sample who was not present. As a result, the adequacy of each replacement, through a comparison of demographic characteristics, was not feasible at baseline.
- Due to limited time and budget, STS and TEAM Girl Malawi did not pilot survey items prior to the operational baseline data collection. Instead, surveys were pretested with a limited number of respondents to assess length of the surveys, Chichewa translations of items and relevance of items for the target population. Without sufficient sample sizes, it was not possible to test the reliability of items before operational baseline data collection, resulting in lower than desired reliability on 2 indicator indices—specifically, IO4.2 and IO4.3. At future evaluation points, additional items may be added to the indices to improve the index reliability measure.
- Although STS trained enumerators on accommodating girls with disabilities during the assessment and provided notes on which girls would require accommodations, only 9 girls (2.40%) used the large-print stimuli accommodation and zero girls used assistive devices such as glasses, magnifiers or hearing aids. This was likely because the number of girls selected into the sample who were identified by TEAM Girl Malawi disability screening partners as needing assistive devices was small, and because the project had not yet distributed devices. Additionally, enumerators only had notes on potential needs for girls who were in the sample and replacement list. Any girls who were assessed but not on either list may not have been offered accommodations that would have been supportive. For comparability, girls will only be provided the assistive devices they used at baseline at subsequent evaluation points.
- Survey responses were sometimes contradictory. For example, the age girls and caregivers self-reported via surveys frequently did not align with the age recorded in the TEAM Girl Malawi enrolment database. In these cases, ages in the database were used for analysis. Additionally, girls' and caregivers' responses to Child Functioning questions were not always consistent. Per FM guidance, analysis of disability prevalence was computed using girls' responses.[40]

Cohort tracking and next evaluation point

To track the same sample beneficiaries from baseline to endline, STS captured the name and unique identifiers of all girls and caregivers. Identifiers will be used to identify girls across evaluation points and to map project monitoring data with evaluation data. STS will rely on TEAM Girl Malawi staff and CBE facilitators to locate sampled girls at the next evaluation point, which will ensure adherence to the longitudinal design of the evaluation. STS will also rely on TEAM Girl Malawi staff to track cohort 1 girls after graduating from CBE to ensure their transition pathway can be measured.[41]

4.5 Qualitative evaluation methodology

Qualitative data collection tools

Table 11 details qualitative data collection tools administered at baseline. A major focus throughout FGDs and KIIs was barriers to girls' education, both in terms of access to school or

[40] For the 11 girls who completed learning assessments but did not respond to the girls' survey, their caregiver responses were used instead.
[41] Attrition buffers were incorporated into sample size calculations to account for girls from the baseline sample who cannot be tracked and assessed in year 3 and year 5 evaluation points. See Annex 6: MEL framework.

CBE, attendance at school or CBE and transition. In order to further understand these barriers, adolescent girls and boys participated in a participatory learning activity (PLA) called 'The Path', which highlighted different impediments at home, on the way to the learning centre and at the learning centre.

Table 11: Qualitative evaluation tools (baseline)

Tool name	Relevant indicator(s)	Who developed the tool?	Was tool piloted?	How were piloting findings acted upon (if applicable)	Was FM feedback provided?
FGD with adolescent girls	O2.1 O2.2 O2.3 IO1.1 IO2.1 IO4.1 IO4.2 IO4.3	STS, Link, TfaC	Yes	Tools were streamlined and questions were cut to reduce length of FGD. Select questions were made optional due to sensitivity for younger respondents. Enumerators were given the option to reverse order to administer PLA first, in case younger respondents needed more 'warm up' to feel comfortable sharing in a group setting.	Yes
FGD with adolescent boys	O2.1 O2.2 O2.3 IO1.1 IO2.1 IO4.1 IO4.2 IO4.3	STS, Link, TfaC	Yes	Tools were streamlined and questions were cut to reduce length. Select questions were made optional due to sensitivity for younger respondents. Enumerators were given option to reverse the order and administer PLA first, in case younger respondents needed more 'warm up' in order to feel comfortable sharing in a group setting.	Yes
FGD with mothers' groups	O2.1 O3 Community IO1.1 IO4.1 IO4.2 IO4.3	STS, Link, TfaC	No	NA	Yes
KII with community leaders	O2.1 O2.2 O2.3 IO1.1	STS, Link, TfaC	No	NA	Yes

Tool name	Relevant indicator(s)	Who developed the tool?	Was tool piloted?	How were piloting findings acted upon (if applicable)	Was FM feedback provided?
	IO1.2 IO1.3 O3 Community IO4.2 IO4.3				
KII with MOEST and MOGCDSW officials (district- and national-level)	O3 System IO5.1 IO5.2 IO5.3	STS, Link, TfaC	No	NA	Yes

Qualitative sample selection and sample sizes

STS employed a combination of purposive and random sampling for the qualitative component of the baseline evaluation. First, a quota of FGDs and KIIs was established and detailed in the evaluation inception report. Link subsequently requested 3 additional FGDs for adolescent girls—for a total of 2 per district—and 3 FGDs with mothers' groups—for a total of one per district.

Table 12 details final subgroups and sample sizes, and Supplemental Table 5 provides qualitative data quotas. TEAM Girl Malawi provided a list of potential subgroups to sample for adolescent girls' FGDs.[42] Appropriate age bands—10 through 14 and 15 through 19—were established for both adolescent girls' and adolescent boys' FGDs.[43] After randomly selecting CBEs for the quantitative sample, STS explored frequencies of beneficiaries within subgroups across the sampled CBEs in each district. Based on the number of beneficiaries within each subgroup from sampled CBEs, STS purposively selected CBE communities and FGD subgroups for each district—one CBE in Lilongwe, 2 in Dedza and one in Mchinji.[44]

After determining the CBE and FGD subgroups per district, STS randomly selected respondents within each subgroup to participate in adolescent girls' and adolescent boys' FGDs. CBE facilitators were provided with the list of randomly selected beneficiaries and asked to recruit as many as were available. All mothers' group members from the respective community were invited to participate in the respective FGD. FGD subgroup sample lists provided by STS ranged in size from 5 to 12 beneficiaries, although turnout was lower in some cases. Data collectors expressed difficulties in mobilising potential respondents to the FGDs on the assigned days. Although TEAM Girl Malawi attempted to mitigate this by providing transport to respondents, turnout was still lower than desired.

KII respondents were selected purposively by TEAM Girl Malawi, who also scheduled all interviews. The MOEST representative from Lilongwe district was unavailable during the week of data collection, as were the national-level MOEST and MOGCDSW representatives. TEAM Girl Malawi staff conducted KIIs with those stakeholders in August 2019.

[42] Based on the programme's marginalisation criteria.
[43] Age bands for the qualitative sample differed from those for the quantitative sample and the transition pathways. Because of the small number of girls and boys in the other age bands, 2 qualitative age bands were created to ensure sufficient FGD respondents.
[44] STS selected 2 CBE communities in Dedza in order to administer FGDs in as many subgroups as possible and to ensure enough respondents per subgroup. Within Dedza, one FGD with adolescent girls and one FGD with mothers' groups were conducted in Kamundi. One FGD with adolescent girls and one FGD with adolescent boys were conducted in Kanyama-Mkomeko. The community leader KII was conducted in Kanyama-Mkomeko.

Table 12: Qualitative sample sizes

Tool (used for which outcome and IO indicator)	Subgroup	Sample size agreed in MEL framework	Actual sample size	Remarks on why there are major differences between anticipated and actual sample sizes (if applicable)
FGD with adolescent girls	Aged 10–14	6–10 respondents per FGD	1 FGD with 5 participants	Mobilization of subgroup respondents was difficult
	Aged 15–19		1 FGD with 9 participants	NA
	Caregivers (pregnant or young mothers)		1 FGD with 8 participants	NA
	Married		1 FGD with 8 participants	NA
	Girls with disabilities		1 FGD with 5 participants	Mobilisation of subgroup respondents was difficult
	Orphaned		1 FGD with 3 participants	Mobilisation of subgroup respondents was difficult
FGD with adolescent boys	Aged 10–14		1 FGD with 5 participants	Mobilisation of subgroup respondents was difficult
	Aged 15–19		2 FGDs, 10 participants total	Mobilisation of subgroup respondents was difficult
FGD with mothers' groups	NA		3 FGDs, 25 participants total	NA
KII with community leaders	NA	3	3 KIIs[45] – one per district; 2 male, 1 female	NA
KII with government officials	MOEST (district-level)	3	3 KIIs[46] – one per district; 2 male, 1 female	NA
	MOEST (national-level)	1	1 KII[47] – CBE National Coordinator; female	NA
	MOGCDSW (district-level)	3	3 KIIs – one per district; 2 male, 1 female	NA
	MOGCDSW (national-level)	1	1 KII[48] – Principal Disability Programmes Officer; 1 male	NA

[45] The secretary for the Mchinji community leader was present and provided comments during KII. Analysis focused on responses provided by the community leader.
[46] TEAM Girl Malawi conducted Lilongwe KII.
[47] TEAM Girl Malawi conducted KII.
[48] TEAM Girl Malawi conducted KII.

Supplemental Table 5: Qualitative evaluation tool quotas

Tool	District[49]			Total
	Lilongwe (Cobayo)	Dedza (Kamundi and Kanyama-Mkomeko)	Mchinji (Kamphata)	
FGD with adolescent girls	2	2	2	6
FGD with adolescent boys	1	1	1	3
FGD with mothers' groups	1	1	1	3
Total FGDs	**4**	**4**	**4**	**12**
KII with community leaders	1	1	1	3
MOEST (district-level)	1	1	1	3
MOEST (national-level)	NA	NA	NA	1
MOGCDSW (district-level)	1	1	1	3
MOGCDSW (national-level)	NA	NA	NA	1
Total KIIs	**3**	**3**	**3**	**11**

Qualitative field researchers

Similar to the selection and hiring process for the quantitative enumerators, STS and CERT worked collaboratively to recruit, hire and train qualitative field researchers for the pretest and operational baseline data collection activities. CERT used qualifications and job descriptions to recruit 2 female field researchers through screenings, oral interviews, reference checks and final review from STS. The field researchers had extensive prior experience with qualitative research, including administering FGDs and KIIs with adolescents and on SRHR and gender-based violence topics. Both were fluent in Chichewa and English.

Before training commenced, the selected field researchers signed contracts with CERT that stipulated their expected roles and professional conduct during training and data collection. Additionally, both field researchers underwent police security clearance checks as required by TEAM Girl Malawi as part of its child safety and protection procedures for all persons working under their projects.

The baseline qualitative researcher training—facilitated by STS with support from CERT, Link and TfaC—took place from 26 to 28 June 2019 in Lilongwe. Training sessions covered the objectives of the qualitative component of the TEAM Girl Malawi baseline study, CP and safeguarding policies and qualitative research practices. It also included an overview and practice of each FGD and KII tool. On the second day of training, the qualitative researchers conducted one pretest of the adolescent girls FGD and one of the adolescent boys FGDs in a non-sampled CBE community outside of Lilongwe. STS provided researchers with constructive feedback following the pretest and used observations and notes from the researchers to update and finalise the 2 tools.

Both field researchers were trained in facilitation and note-taking to enable them to rotate roles during the data collection.

Qualitative data collection

[49] Qualitative data was collected in 2 different CBEs in Dedza and in one CBE each for Lilongwe and Mchinji.

Qualitative data collection took place from 29 June to 5 July 2019, one week prior to the quantitative data collection.[50] STS drafted a schedule of CBE visits and the timeslots for each qualitative activity. Each CBE visit spanned 2 days, and no more than 3 FGDs were scheduled per day. Link distributed the proposed CBE visit schedule and selected respondents to CBE facilitators, who were responsible for scheduling and respondent mobilisation. Field researchers contacted TEAM Girl Malawi district staff one day prior to their visit to reconfirm the schedule of activities and ensure respondent participation. Due to election-related protests and community events, the researchers had to revise their visit schedule during the week of data collection. However, because the distances between selected CBEs were relatively close, these disruptions did not necessitate an extension of the data collection period.

FGD tools and community leader KII tools were translated and administered in Chichewa, while government official KIIs were administered in English. The field researchers alternated serving as facilitators and note-takers for FGDs and KIIs during the data collection. All FGDs and KIIs were recorded.[51] Researchers took detailed field notes and reflections during the activities. They were required to submit a set of documents—a daily debrief form, FGD seating charts, FGD registration forms and photos of 'The Path' activity outputs—via STS' secured, password-protected server at the end of each night. Researchers were requested to complete an expanded notes template in Microsoft Word in English for each FGD and KII, in which findings, direct quotes and reflections were described and supplemented by the audio-recordings. Although STS requested that these were submitted each night, researchers ultimately completed these within a 2 to 3-day time period given the rigor of the data collection schedule and the quantity of qualitative data collected each day.

STS reviewed documents daily for completeness and outstanding questions, concerns or clarifications. STS and the qualitative field researchers communicated during data collection by WhatsApp, following up with questions about the data, quotas and logistical challenges that may have been encountered.

Qualitative data handling and analysis

Field researchers managed transcription and translation according to STS guidance. The notetaker took handwritten field notes during FGDs and KIIs.[52] Utilising the handwritten field notes and the audio-recording as references, the notetaker and facilitator collaboratively completed the expanded notes template in English. The most pertinent quotes were also transcribed verbatim in both Chichewa or English and included in the expanded field notes. Field researchers did not complete verbatim transcripts and translations. However, they reviewed and cross-checked each others' expanded field notes and translations of key quotes from local language to English to ensure quality and accuracy.

Qualitative researchers uploaded all data—including audio-recordings and expanded field notes—to STS' secured, password-protected server. After all raw data were confirmed as successfully uploaded to the server, CERT ensured that audio-files were deleted from the

[50] Qualitative data collection took place the week of the quantitative enumerator training for several reasons. First, it ensured that an STS staff member was in-country during the data collection to provide support and troubleshooting. Also, it alleviated the burden on respondents, who were, as a result, not required to participate in qualitative and quantitative activities during the same time period. Last, the sequencing was based on budget. Because there was not sufficient budget to support two STS team members in-country, it allowed for the same STS team member to deliver training for the qualitative and quantitative activities across 2 weeks.
[51] If permission was granted by respondents.
[52] This included quotes, key points and themes that emerged for each question, non-verbal activity or body language, as well as any big ideas, thoughts or take-aways from the note-taker.

recording device. All qualitative materials—including, notebooks, completed FGD and KII guides and FGD registrations—were returned to CERT after the completion of data collection.

Finalised expanded field notes were imported into NVivo 12, a qualitative data analysis software package, in order to systematically code and analyse the data. All coding was completed by a single user. The qualitative data analysis methodology incorporated an iterative approach and included content analysis and constant comparison of narrative data to identify and validate emerging themes. While a preliminary codebook was developed based on the LNGB baseline report template as well as the TEAM Girl Malawi baseline study core research themes and key concepts, additional codes that emerged during the data analysis were incorporated and the codebook updated as needed. STS examined qualitative data and emergent themes within the broader context of the quantitative results and indicators. Relevant findings were woven into the report as appropriate to help provide additional insights and understanding into the TEAM Girl Malawi baseline evaluation results, analyses and recommendations. While observations by researchers are included in the qualitative analysis where relevant, reflections and recommendations are clearly distinguished from the raw data and findings.

Challenges in baseline qualitative data collection, handling and analysis and limitations of the qualitative aspects of the evaluation design

STS and TEAM Girl Malawi faced several key challenges during the qualitative data collection and analysis:

- The number of FGDs conducted at baseline was limited due to logistical and budget constraints, as well as distribution of subgroup populations across and within CBEs. While the qualitative sample would ideally include at least 2 FGDs per targeted subgroup, homogenous FGDs for each subgroup of adolescent girls was prioritised to ensure that the experiences and voices of the most marginalised girls in the project were captured. Qualitative findings from adolescent girls by subgroup should be understood as findings from a single focus group.
- Due to time, budget and logistical constraints, detailed field notes were utilised in place of fully translated transcriptions. Expanded field notes produced by the note-taker enabled a quicker turnaround that was less labour intensive and fit within the budget constraints the baseline evaluation. However, the discussions, reflections and insights from FGDs and KIIs may be limited due to a lack of full transcriptions and translations.
- Audio-recordings of ‘The Path’ activity were not available due to the way activity was administered. Specifically, the facilitator and note-taker separated the participants into 2 groups to facilitate 2 smaller discussions and did not record the small group discussions. Instead, photos of the completed activity, including the flipchart sheets with notes highlighting the key barriers and solutions identified by participants for each category served as the raw data for analysis. When conducting this activity in the future, 2 audio-recording devices will be made available to ensure recordings are captured to serve as additional reference for field notes.
- Lack of responsiveness and unwillingness to engage in FGDs, especially by girls aged 10 to 14, was noted as a challenge by qualitative field researchers. Although icebreakers were used to help put participants at ease and FGD included one PLA activity, additional PLAs may be considered to encourage more participation in the future. Furthermore, limiting age groups for FGDs to within 2 or 3 years may also be considered.
- Three government official KIIs were conducted directly by TEAM Girl Malawi staff because officials were unavailable during the qualitative data collection period. These KIIs are noted in Table 12. Because the KII facilitator for was not directly trained by STS, there

may have been differences in the way that questions were asked or recorded. Additionally, because the facilitator was a TEAM Girl Malawi staff member, bias may have been introduced during the interview and when taking notes.

- Qualitative data collection was disrupted due to election-related political unrest. Specifically, qualitative researchers were, in some instances, unable to reach communities as intended because of roadblocks and strikes. These disruptions also impacted the availability of respondents, particularly government officials. To mitigate these challenges, TEAM Girl Malawi staff and qualitative field researchers communicated frequently to make modifications to appointments and to reschedule with respondents as needed.

5. Key characteristic subgroups and barriers of baseline samples

The following section examines the main characteristics of the subgroups of interest and the barriers to learning and transition that they face. This section also examines the intersection between the main barriers and characteristics to help determine how appropriate the TEAM Girl Malawi project activities are for these subgroups and if the ToC is accurate. Barriers were drawn from the qualitative study, and STS used surveys to quantify barrier prevalence.

5.1 Educational marginalisation

TEAM Girl Malawi identified characteristic subgroups, which are a critical part of girls' enrolment marginalisation criteria. These 8 subgroups are detailed in Table 13, and the items used to construct marginalisation subgroups are included in Supplemental Table 34. One of TEAM Girl Malawi's subgroups—persons with albinism—was not included in analysis due to low prevalence in the beneficiary population.[53] One additional characteristic subgroup was identified during analysis—girls who are married and caregivers—and is included in analysis because of the high overlap between these 2 component subgroups.[54] The compound subgroup will be used for analyses throughout this report. Further, the high poverty subgroup will not be reported throughout this report, as all respondents in the sample reported high poverty. Instead, analyses looking at the relative impact of the barriers and marginalisation uses the extent of hunger experienced by the family as a proxy for extreme poverty within a high poverty population.

Table 13: Characteristic subgroups

Characteristic	Description	Proportion of sample with this characteristic
High poverty	Girl is over poverty threshold	100.00%
High chore burden	Girl has high number of chore hours (6 or more per day)	47.88%
Caregiver	Girl is primary caregiver for her own or other children	46.30%
Girls with disabilities	Girl has functional difficulty	34.39%
Orphaned	Girl lost one or both parents	26.19%
Married	Girl is, was, or is about to be married	20.90%

[53] According to Link, only one girl enrolled in the programme is a person with albinism. The programme credits low enrolment of girls with albinism to the high levels of social stigma and insecurity experienced by these girls, which was heightened by the lead up to the elections in May 2019. The programme is examining safeguarding approaches to including girls with albinism in future cohorts.
[54] There were statistically significant correlations between being married and being a caregiver. Specifically, 94.94% of girls who are married are also caregivers and 42.86% of caregivers are also married. See Supplemental Table 45.

Characteristic	Description	Proportion of sample with this characteristic
Married and caregiver	Girl is/was/about to be both married and a caregiver	19.84%
Head of household	Girl is head of household	2.91%
Source: N = 378	TEAM Girl Malawi enrolment database, girls survey, household survey	

To categorise sample respondents into characteristic subgroups, STS merged the TEAM Girl Malawi enrolment database with baseline survey respondents. Then, STS used a specified set of items to determine whether a girl is part of a characteristic subgroup. Findings indicate that 26.19% of girls in the sample are orphaned, 47.88% have a high chore burden, and 2.91% are heads of household. About one in 5 girls in the sample are, were or are about to be married. Nearly half are the primary caregivers for their own or others' children. The overlap between these 2 categories—girls who are married and girls who are caregivers—shows that 19.84% girls in the sample are married caregivers.

Key barriers to learning and transition are listed in Table 14. Specific items used to categorise girls into barrier subgroups are detailed in Supplemental Table 35. To populate these barriers, STS used a mixed-methods approach. First, STS analysed baseline qualitative data from FGDs with adolescent girls and boys to identify the key barriers mentioned by beneficiaries. The key barriers identified through qualitative results were cross referenced with survey items to determine which barriers could be quantified in the sample population.[55] STS also cross-referenced barriers with TEAM Girl Malawi's needs assessment, completed in 2018, to ensure that those barriers most frequently identified in the population were included in the baseline analysis.

Table 14: Barriers

Barrier	Barrier description	Proportion of sample affected by this barrier
School cost	School cost	88.36%
Food insecurity or hunger	Household experiences hunger 10 or more days per year	62.42%
Menstruation	Girl has started menstruation, has low self-efficacy during menstruation and limited access to support and supplies to manage menstruation	33.07%
School safety	School or going to school is not safe	27.25%
Parent support	Lack of parental support for school	18.52%
Bullying	Girl is bullied or fears being bullied by teacher or peers	7.67%
Source: N = 378	FGDs with adolescent girls, TEAM Girl Malawi enrolment database, girls survey, household survey	

Findings indicate that school cost is the most frequently experienced barrier—88.36% of girls experienced this barrier at baseline. Food insecurity or hunger and menstruation were also

[55] All of the main barriers identified by girls and boys in qualitative data could quantified using the enrolment database and baseline surveys.

mentioned by girls—62.42% of girls experienced hunger 10 or more days per year and 33.07% of girls experienced menstruation as a barrier.[56]

5.2 Intersection between key characteristics subgroups and barriers

The intersections between characteristic subgroups and barriers are presented in Table 15 and Supplemental Table 6.[57] Across all subgroups, school cost was a frequently reported barrier. Given that all girls in the sample had high levels of poverty, STS expected this finding. Results were similar for the intersection of subgroups and food insecurity or hunger. Over half of respondents in all subgroups, except head of household, reported food insecurity or hunger as a barrier. About 3 in 5 of girls who reported bullying as a barrier had a functional disability, and 61.43% of the girls who reported a lack of parental support as a barrier were in the high chore burden subgroup.

Qualitative findings did not provide any additional explanatory details regarding the prevalence of barriers faced by girls in specific subgroups.

[56] A menstruation barrier index was constructed to identify girls experiencing the barrier. See Supplemental Table 35 for a list of items.
[57] Statistical (chi-squared) tests are not included on relationships between marginalisation characteristics and barriers, as the sample was not powered to be large enough to make generalisations within subgroups.

Table 15: Key barriers to education by characteristic subgroups

		Characteristics						
		High chore burden (n=181)	Caregiver (n=175)	Girls with disabilities (n=130)	Orphaned (n=99)	Married (n=79)	Married and caregiver (n=75)	Head of household (n=11)
Barriers	School cost (n=334)	48.80% of girls who reported school cost as a barrier were in the high chore burden subgroup.	43.11% of girls who reported school cost as a barrier were in the caregiver subgroup.	35.63% of girls who reported school cost as a barrier were in the girls with disabilities subgroup.	28.74% of girls who reported school cost as a barrier were in the orphaned subgroup.	20.06% of girls who reported school cost as a barrier were in the married subgroup.	18.86% of girls who reported school cost as a barrier were in the married and caregiver subgroup.	2.69% of girls who reported school cost as a barrier were in the head of household subgroup.
	Food insecurity or hunger (n=236)	51.27% of girls who reported food insecurity or hunger as a barrier were in the high chore burden subgroup.	44.92% of girls who reported food insecurity or hunger as a barrier were in the caregiver subgroup.	33.90% of girls who reported food insecurity or hunger as a barrier were in the girls with disabilities subgroup.	29.24% of girls who reported food insecurity or hunger as a barrier were in the orphaned subgroup.	22.46% of girls who reported food insecurity or hunger as a barrier were in the married subgroup.	20.76% of girls who reported food insecurity or hunger as a barrier were in the married and caregiver subgroup.	1.27% of girls who reported food insecurity or hunger as a barrier were in the head of household subgroup.
	Menstruation (n=125)	53.60% of girls who reported menstruation as a barrier were in the high chore burden subgroup.	50.40% of girls who reported menstruation as a barrier were in the caregiver subgroup.	48.31% of girls who reported menstruation as a barrier were in the girls with disabilities subgroup.	26.40% of girls who reported menstruation as a barrier were in the orphaned subgroup.	24.00% of girls who reported menstruation as a barrier were in the married subgroup.	22.40% of girls who reported menstruation as a barrier were in the married and caregiver subgroup.	4.00% of girls who reported menstruation as a barrier were in the head of household subgroup.
	School safety (n=103)	48.54% of girls who reported lack of school safety as a barrier were in the high chore burden subgroup.	40.78% of girls who reported lack of school safety as a barrier were in the caregiver subgroup.	40.78% of girls who reported lack of school safety as a barrier were in the girls with disabilities subgroup.	27.18% of girls who reported lack of school safety as a barrier were in the orphaned subgroup.	18.45% of girls who reported lack of school safety as a barrier were in the married subgroup.	18.45% of girls who reported lack of school safety as a barrier were in the married and caregiver subgroup.	1.94% of girls who reported lack of school safety as a barrier were in the head of household subgroup.
	Lack of parental	61.43% of the girls who	41.43% of girls who reported	47.14% of the girls who	27.14% of girls who reported	24.29% of girls who reported	22.86% of girls who reported	0.00% of the girls who

		Characteristics						
		High chore burden (n=181)	Caregiver (n=175)	Girls with disabilities (n=130)	Orphaned (n=99)	Married (n=79)	Married and caregiver (n=75)	Head of household (n=11)
	support (n=70)	reported lack of parental support as a barrier were in the high chore burden subgroup.	lack of parental support as a barrier were in the caregiver subgroup.	reported lack of parental support as a barrier were in the girls with disabilities subgroup.	lack of parental support as a barrier were in the orphaned subgroup.	lack of parental support as a barrier were in the married subgroup.	lack of parental support as a barrier were in the married and caregiver subgroup.	reported lack of parental support as a barrier were in the head of household subgroup.
	Bullying (n=29)	24.14% of girls who reported bullying as a barrier were in the high chore burden subgroup.	31.03% of girls who reported bullying as a barrier were in the caregiver subgroup.	62.07% of girls who reported bullying as a barrier were in the girls with disabilities subgroup.	24.14% of girls who reported bullying as a barrier were in the orphaned subgroup.	13.79% of girls who reported bullying as a barrier were in the married subgroup.	13.79% of girls who reported bullying as a barrier were in the married and caregiver subgroup.	0.00% of girls who reported bullying as a barrier were in the head of household subgroup.

Supplemental Table 6: Key barriers to education by characteristic subgroups (characteristic as independent variable)

		Characteristics						
		High chore burden (n=181)	Caregiver (n=175)	Girls with disabilities (n=130)	Orphaned (n=99)	Married (n=79)	Married and caregiver (n=75)	Head of household (n=11)
Barriers	School cost (n=334)	90.06% of girls in the high chore burden subgroup reported school cost as barrier.	82.29% of girls in the caregiver subgroup reported school cost as barrier.	91.54% of girls in the girls with disabilities subgroup reported school cost as barrier.	96.97% of girls in the orphaned subgroup reported school cost as barrier.	84.81% of girls in the married subgroup reported school cost as barrier.	84.00% of girls in the married and caregiver subgroup reported school cost as barrier.	81.82% of girls in the head of household subgroup reported school cost as barrier.
	Food insecurity or hunger (n=236)	66.85% of girls in the high chore burden subgroup reported food insecurity or hunger as a barrier.	60.57% of girls in the caregiver subgroup reported food insecurity or hunger as a barrier.	61.54% of girls in the girls with disabilities subgroup reported food insecurity or hunger as a barrier.	69.70% of girls in the orphaned subgroup reported food insecurity or hunger as a barrier.	67.09% of girls in the married subgroup reported food insecurity or hunger as a barrier.	65.33% of girls in the married and caregiver subgroup reported food insecurity or hunger as a barrier.	27.27% of girls in the head of household subgroup reported food insecurity or hunger as a barrier.

		Characteristics						
	Menstruation (n=125)	37.02% of girls in the high chore burden subgroup reported menstruation as a barrier.	36.00% of girls in the caregiver subgroup reported menstruation as a barrier.	43.85% of girls in the girls with disabilities subgroup reported menstruation as a barrier.	33.33% of girls in the orphaned subgroup reported menstruation as a barrier.	37.97% of girls in the married subgroup reported menstruation as a barrier.	37.33% of girls in the married and caregiver subgroup reported menstruation as a barrier.	45.45% of girls in the head of household subgroup reported menstruation as a barrier.
	School safety (n=103)	27.62% of girls in the high chore burden subgroup reported lack of school safety as a barrier.	24.00% of girls in the caregiver subgroup reported lack of school safety as a barrier.	32.31% of girls in the girls with disabilities subgroup reported lack of school safety as a barrier.	28.28% of girls in the orphaned subgroup reported lack of school safety as a barrier.	24.05% of girls in the married subgroup reported lack of school safety as a barrier.	19.84% of girls in the married and caregiver subgroup reported lack of school safety as a barrier.	18.18% of girls in the head of household subgroup reported lack of school safety as a barrier.
	Lack of parental support (n=70)	23.76% of girls in the high chore burden subgroup reported lack of parental support.	16.57% of girls in the caregiver subgroup reported lack of parental support.	25.38% of girls in the girls with disabilities subgroup reported lack of parental support.	19.19% of girls in the orphaned subgroup reported lack of parental support.	21.52% of girls in the married subgroup reported lack of parental support.	21.33% of girls in the married and caregiver subgroup reported lack of parental support.	0.00% of girls in the head of household subgroup reported lack of parental support.
	Bullying (n=29)	7.07% of girls in the high chore burden subgroup reported bullying.	5.14% of girls in the caregiver subgroup reported bullying.	13.85% of girls in the girls with disabilities subgroup reported bullying.	7.07% of girls in the orphaned subgroup reported bullying.	5.06% of girls in the married subgroup reported bullying.	5.33% of girls in the married and caregiver subgroup reported bullying.	0.00% of girls in the head of household subgroup reported bullying.

5.3 Appropriateness of project activities to the characteristic subgroups and barriers identified

1. *Are there any additional characteristic subgroups revealed through the baseline data collection that may be at risk of educational marginalisation that are not considered in project intervention planning?*

Baseline data did not reveal any unanticipated characteristic subgroups that are not considered in intervention planning.

2. *Do the most prevalent barriers identified by the analysis conducted by the EE correspond with the project's ToC? Or are there any additional barriers to learning or transition that were not considered in project intervention planning?*

The most prevalent social, economic and educational barriers uncovered through the baseline are being considered in TEAM Girl Malawi intervention planning. These include support for girls' SRHR—specifically menstrual health— through Girls' Clubs, financial support through micro-loans for households with poverty or food insecurity and system-level support for families. The project should ensure that school safety—both on the way to CBE and at CBE—is prioritised, as are sensitivity and awareness-building activities for girls' caregivers, particularly those who have disabilities. Further, improving caregiver support for girls' education should be emphasised in trainings.

Because girls' frequently experience food insecurity and hunger as barriers, the project may consider incorporating food assistance or feeding at CBEs as part of its intervention. Alternatively, if feeding programmes are out of scope for TEAM Girl Malawi, the project may consider partnering with other organisations or programmes to try to address this barrier.

Not all the barriers listed in the project's ToC were evidenced through the baseline data—namely, the educational marginalisation barriers. This is primarily due to the respondents that participated in the baseline – specifically, out-of-school girls and their family members instead of stakeholders and beneficiaries in the formal school system. Nevertheless, findings from the baseline do not explicitly disprove any of the educational barriers that were identified by the project through the GESI, and several were substantiated through qualitative findings (see Supplemental Table 23). TEAM Girl Malawi may want to revisit assumed educational barriers through monitoring to ensure that they continue to be applicable to the beneficiary population and communities.

3. *Do the project interventions address the key barriers for the key characteristic subgroups?*

The project interventions appear to address key barriers for key characteristic subgroups. School safety—including bullying—should be monitored routinely for girls with disabilities. Parental engagement in and support for girls' education should be a focus for girls with high chore burdens and girls with disabilities.

The project should also ensure that girls who are caregivers are provided with appropriate childcare support to enable their participation in and regular attendance at project activities. The project should also ensure that girls with high chore burdens or who are currently in paid employment (5.18% of the sample; see Table 20) are linked with financial support to enable their participation. TEAM Girl Malawi should further explore school fees and fees that would be levied on girls seeking to transition back into the formal school system.

There is a relatively high prevalence of girls with functional difficulties, according to results from the Child Functioning module. Recognising that these girls face intersectional challenges that are generally understood, such as accessibility at learning centers, and that were elucidated in the

baseline, such as bullying, the project should ensure that specific interventions to create equal access and inclusive education, as well as monitoring plans for attendance and learning.

4. Do the assumptions in the theory of change hold true?

Assumptions in the project's ToC regarding subgroups and barriers appear to hold true. These assumptions will be revisited at the next evaluation point.

Project to complete

- The project should respond to the external evaluators' comments on the above questions. In particular the project should respond to:
    - Why the projects theory of change may not correspond with some of the key barriers or characteristic subgroups identified.
    - Whether the project plans to review some aspects of their Theory of change in light of these findings.

TEAM Girl Malawi understands the fundamental impact of extreme poverty as the intrinsic and pervasive barrier to education. This was identified prior to beneficiary selection and confirmed through the enrolment and baseline findings. As such, consideration and monitoring will be given to attrition rates that could be attributed to this barrier and all affiliated impacts. The microloan, business training and vocational training aspects of the intervention are designed to alleviate economic burden on the households of learners to allow them to attend CBEs and Girls' Clubs. However, it should be considered that these interventions will not have an immediate and universal impact. Therefore, the project may still see some attrition related to such issues, particularly where chores, agricultural activity and household mean are prioritised and take precedence over education.

Many girls noted that food insecurity and hunger as frequently experienced issues. While direct feeding projects are beyond the scope and resource of the project, TEAM Girl Malawi will look into the potential of collaborating with other projects to meet such needs—especially in food insecure times of year.

Mothers, caregivers, heads of households, married and pregnant young girls are extremely prevalent in cohort 1. TEAM Girl Malawi is working with these beneficiaries and their communities toward measures and support mechanisms that will best serve their ability to attend and get beyond any such issues which would perpetuate as barriers to education. The best way the project can do this is to make learning environments and approaches as inclusive and accommodating as possible, but it must also work with the households, communities and school and statutory services for sustainable change and impact.[58]

School safety—specifically in travelling to and from classes and from a safeguarding perspective—has been a focus of the project more recently. Specifically, it has been considering distance to classes, bullying and stigma, access for learners with disabilities, time of day and daylight, foliage and state of routes at different times of year. Some detail and context specific

[58] In pre-CBE, TEAM Girl Malawi delivered sessions on co-operation, respect and ground rules. Girls Club sessions will incorporate bullying themes by addressing inclusivity within all workshops. TEAM Girl Malawi is currently working on a set of guidelines to explicitly address stigmas and bullying.

scenarios as well as—and more importantly—beneficiary and wider stakeholder feedback will help TEAM Girl Malawi learn and adapt its approach.

We are pleased to read that our assumptions in the ToC hold true and will continue to refer to causality and pathways identified. We must however learn and adapt accordingly to challenges as we respond to specific and challenging vulnerabilities and entrenched barriers.

6. Outcome findings

Baseline results for the following TEAM Girl Malawi outcomes are presented in this section:

- O1: Number of highly marginalised girls supported by GEC with improved learning outcomes[59]
- O2: Number of marginalised girls who have transitioned through key stages of education, training or employment
- O3: Project can demonstrate that the changes it has brought about which increase learning and transition through education cycles are sustainable

6.1 Learning outcomes

TEAM Girl Malawi's first outcome is improved learning outcomes. This section will present findings on the following indicators:

- O1.1: Number of highly marginalised girls supported by GEC with improved literacy outcomes
- O1.2: Number of highly marginalised girls supported by GEC with improved numeracy outcomes

Baseline findings for the third learning outcome—O1.3 Number of highly marginalised girls supported by GEC with improved life skills outcomes—are detailed in Section 7.2.

Project to complete

- Please outline the learning levels girls have started with and what level you are aiming girls to reach by the next evaluation point and, if applicable, once they complete the full learning intervention. This should reflect any differences in ambition depending on the intervention pathway of characteristic subgroup.
- If benchmarking was used, provide a summary of what levels or grades you used for benchmarking and why.

Beneficiaries were eligible if they did not have functional literacy and numeracy levels or if they had been out of school and between the ages of 10 and 19. Eligibility was determined through the Aser test: candidates were eligible for the project if they were unable to complete the highest-level literacy and numeracy tasks. The CBE curriculum is designed to cover standards 1, 2, 3 and 4 in a 2-year period. The girls would then be at a level of literacy, numeracy and knowledge in key subjects that would see them re-enter school at standard 5—should they wish to take that transitional pathway—or have them at a sufficient level to attend business or vocational training.

[59] Baseline results for O1.3 Number of highly marginalised girls supported by GEC with improved life skills outcomes are presented in section 7.2.

Results were benchmarked against MOEST standards 1 through 3, which was also the source of the literacy assessments.

The second evaluation point (2021) will re-assess girls from cohort 1 after they complete their full learning intervention. The project appreciates that there will be a range of abilities, ages and challenges for subgroups. Additionally, the curriculum may suit some abilities better than others. The project aims to reassess through the early stages of cohort 1 to determine how best to adapt its approach if needed.

Headline results

Girls' baseline literacy findings are presented in Table 16. Apart from the letter name identification and listening comprehension subtasks, a majority of girls are categorised as 'non-learners'—meaning that they received zero scores on a given subtask. Specifically, 79.63% of girls did not read a single word in the oral reading fluency passage, and 82.54% of girls did not answer a single reading comprehension question correctly.[60] Zero score proportions were also high on the foundational skills subtasks, including initial sound identification and syllable identification—69.84% and 64.81%, respectively. A smaller proportion of girls received zero scores on the letter name identification subtask (48.41%). Reading of sight words also proved challenging for girls; 71.43% did not correctly read a single item on the familiar word reading subtask. Girls did, however, perform better on the listening comprehension subtask, where the largest proportion of girls—44.71%—scored as 'established learners'.

There do not appear to be ceiling effects at baseline. On most subtasks, no more than 10% of girls scored as 'proficient learners'. However, on listening comprehension, 20.37% of girls scored as proficient learners. Because listening comprehension tests oral vocabulary and not reading, it is expected that girls would perform better on this subtask than on reading subtasks. On no other subtasks did more than one in 10 girls score as a proficient learner.

Given these findings, the project appears to have accurately targeted girls without functional literacy. Indicator O1.1 will measure improved literacy outcomes of girls participating in the project, and due to the low literacy levels at baseline and the low risk of fluency-rate ceiling effects on timed subtasks and percentage-correct ceiling effects on untimed subtasks, there is substantial room for literacy improvement during girls' 2 years of CBE.

Girls' baseline numeracy findings are presented in Table 17. Girls appeared to have stronger performance in mathematics than on literacy subtasks. Overall, there is a wider distribution of performance across the learner categories, and, in general, fewer girls did not answer a single item correctly on a subtask. About one in 4 girls (26.19%) scored as proficient learners on the number recognition subtask, while one in 10 (10.58%) received zero scores. The largest proportion of girls (35.98%) scored as established learners on the quantity discrimination subtask, with similar proportions scoring as emergent learners and non-learners (27.25% and 27.51%, respectively). More than half of girls received zero scores on higher order addition and subtraction subtasks—53.17% on addition level 2 and 56.08% on subtraction level 2. Although 22.22% of girls received zero scores on word problems, 12.43% scored as proficient learners.

Although performance on mathematics subtasks was stronger than on literacy subtasks, ceiling effects do not appear to be a concern when examining the fluency rates for timed subtasks and the percentage correct scores for untimed subtasks. Although about one-quarter of girls were

[60] Reading comprehension zero scores are comprised of girls who were not given the opportunity to answer any questions due to receiving a zero score on the oral reading fluency subtask and girls who were asked comprehension questions but did not answer any correctly.

categorised as proficient learners on one numeracy subtask, they will have the ability to increase their speed on this timed task at the next evaluation point. No more than 15% of girls were proficient learners on any of the other mathematics subtasks.

Given the distribution in performance, the project should consider taking a differentiated approach to mathematics teaching. Indicator O1.2 will measure improved numeracy outcomes of girls participating in the project, and given the low risk of ceiling effects, there is substantial room for numeracy improvement during girls' 2 years of CBE.

Table 16: Foundational literacy gaps

Categories	Subtask 1 Initial sound identification	Subtask 2 Letter name identification	Subtask 3 Syllable identification	Subtask 4 Familiar word reading	Subtask 5 Oral reading fluency	Subtask 6 Reading comprehension	Subtask 7 Listening comprehension
Non-learner 0%	69.84%	48.41%	64.81%	71.43%	79.63%	82.54%	4.76%
Emergent learner 1–40%	28.57%	36.24%	19.58%	10.85%	17.99%	5.29%	30.16%
Established learner 41–80%	1.59%	13.76%	9.52%	7.94%	2.38%	9.52%	44.71%
Proficient learner 81–100%	0.00%	1.59%	6.08%	9.79%	0.00%	2.65%	20.37%
Source: N=378	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**

Table 17: Foundational numeracy skills

Categories	Subtask 1 Number recognition	Subtask 2 Quantity Discrimination	Subtask 3 Missing Numbers	Subtask 4a Addition (1)	Subtask 4b Addition (2)	Subtask 5a Subtraction (1)	Subtask 5b Subtraction (2)	Subtask 6 Word problems
Non-learner 0%	10.58%	27.51%	33.86%	28.31%	53.17%	36.24%	56.08%	22.22%
Emergent learner 1–40%	29.63%	27.25%	54.50%	24.60%	31.22%	25.40%	29.63%	31.22%
Established learner 41–80%	33.60%	35.98%	11.11%	32.54%	12.17%	28.04%	10.85%	34.13%
Proficient learner 81–100%	26.19%	9.26%	0.53%	14.55%	3.44%	10.32%	3.44%	12.43%
Source: N=378	100%	100%	100%	100%	100%	100%	100%	100%

6.2 Characteristic subgroup analysis of the learning outcome

Literacy and numeracy aggregate scores by subgroup and barrier are presented in

Table 18. Overall, there was no statistically significant difference in literacy or numeracy performance by district. By age, girls aged 12–16 and girls aged 17–19 both outperformed girls aged 10–11. Girls who had functional difficulties walking, had functional difficulties communicating or reported bullying as a barrier all had statistically significantly lower literacy and numeracy aggregate scores than other girls. Girls in the married and caregiver, orphaned and high chore burden subgroups had statistically higher literacy and numeracy scores than other girls. When interpreting these results, it is important to keep in mind that girls who are married and caregivers, who are orphans or who have a high chore burden were also more likely to be in the older age groups. In other words, these subgroups overlap with the age groups in such a way that it is not possible to distinguish the effects on learning outcomes of age versus a girls' status in these 3 subgroups.[61]

Table 18: Learning scores by key characteristic subgroups and barriers

Subgroup	Average literacy score (aggregate)	Average numeracy score (aggregate)
All girls	**18.31**	**32.23**
Lilongwe	18.62	30.91
Dedza	18.78	31.81
Mchinji	17.24	33.93
Age 10–11	9.48	10.99
Age 12–16^	19.00	35.60
Age 17–19^	22.05	37.60
Disability subgroup		
Seeing	16.01	33.72
Hearing	9.05	28.91
Walking*	11.59	17.44
Self-care	7.79	16.67
Communication*	1.27	5.83
Learning, remembering and concentrating	15.37	28.39
Accepting change, controlling behaviour and making friends	14.79	27.85
Mental health (anxiety and depression)	18.62	32.16
Subgroup		
Married and caregiver*	24.32	39.62
Orphaned*	20.72	36.60
Head of household	20.51	43.90
High poverty	19.56	36.43
High chore burden*	16.35	30.21
Girls with disabilities	24.32	39.62
Barrier		
Bullying*	11.21	21.29
School cost	18.04	32.70
Parent support	19.29	32.80
Menstruation	18.70	35.01

[61] This same consideration applies to these subgroups throughout the remainder of the baseline report.

Subgroup	Average literacy score (aggregate)	Average numeracy score (aggregate)
Food insecurity or hunger	18.45	32.79
School safety	18.38	34.18

Note: One asterisk (*) denotes differences between subgroups and all girls that are statistically significant at $p < 0.05$. One caret (^) denotes that the difference between the group and age 10–11 group is statistically significant at $p < 0.05$.

To understand the relationships between different levels of the TEAM Girl Malawi ToC, average literacy and numeracy scores are presented by IO indicator scores in Supplemental Table 7. First, girls' scores on the index were grouped into 2 categories: high and low (see footnotes below and Section 7). Then, the relationships between learning outcomes and group were examined for IO2.1, IO2.2, IO4.2, IO4.3 and O1.3 (see Section 7 for calculation of high and low groups and definitions of each index; see The project's initial response **to the findings of the baseline report**

The project will take time to further analyse the report and its findings, conclusion and recommendations before strategizing its responses – both in terms of activities and guidelines, specifically where the greatest needs are identified, and respective impact could be felt. This process will also consider available resources. The projects internal monitoring and adaptive management approach will add to the knowledge acquired through the baseline; working with a community led approach, we expect to integrate our knowledge and learning into agreed practices - responding to resolve key issues, both at a community level, and a project wide level.

The project is pleased that the baseline report recognises the targeting approach to beneficiary selection. We believe this vindicates the efforts and methods to reach the specified subgroups defined by the barriers they faced to education. As mentioned, the slightly lower representation of albinism could be down to a number of factors including the timing of the activities – juts before 2019 elections when tensions are higher and negative instances targeting persons with albinism tend to increase.

Reflecting on the ToC – the project remains satisfied that the assumptions which underpin the expected causal linkages are still relevant. However, with more nuanced and detailed knowledge of the specific community contexts and the circumstances and characteristics of the girls – we do expect to further challenge and detail some of these assumptions, however. Having said that – that - the lack of statistical significance (due to sample sizes) of the some of the findings relating to sub-groups, mean that although we can confidently infer from the findings, further research and analysis throughout cohort 1 should inform any major program shifts. However, the baseline has highlighted many issues and focussed our attention on more pertinent and complex matters.

The main findings involving barriers to learning and transition were largely expected – the reference specifically to school costs and levels of poverty. The recognised prominence of this was demonstrated by the fact that extreme poverty was consistently ranked as the number one barrier in the vulnerability criteria used for targeting. Extreme poverty was considered a barrier for 100% of selected beneficiaries - it is underlying and interlinked with many other conditions and characteristics. Further analysis will support our approaches to understand specific barriers encountered by beneficiaries facing different issues due to localised circumstances, or multiple barriers which can exacerbate or entrench an individual's inability to regularly attend school or be able to commit time in general.

The other more specific issues raised were around food security and hunger – the prominence of this, although not unexpected due to the inescapable correlation with poverty, was higher than had been initially planned for. It is recognised in the contextual understanding that Malawi often faces basic food shortages due to irregular rains/drought coupled with the domestic reliance on subsistence crops and few staple products such as maize and sorghum. However, the ToC did not consider the direct impact of food security and hunger as a barrier, but rather aimed to boost the economic circumstances of beneficiary households through vocational/business training and loan groups. It seems however that the impact of food security could be felt more acutely, and despite the initial forecast of a bumper harvest in 2019, some predictions indicate a shortage of food towards the planting season and traditionally lean time of year. As mentioned previously, it is considered outside of the direct remit and resource capabilities of the project to monitor malnutrition, and/or provide food as a core component (not least as it is quite unsustainable) - however it recognised that there must be efforts made to address such a prominent barrier before waiting for the anticipated impact of the economic benefits to be felt (through transition pathways and households support to parents/guardians).

Options being considered are to look into the potential of partnerships with government agencies/NGOs, and projects already working with food distribution and security. However, this would likely be on and bespoke basis – not blanket coverage, and time specific (to more food insecure intervals), not continuous throughout the project. It will be important to do this pre-emptively, and through further on-the-ground analysis, as the increase of absenteeism could be sudden and severe. The relative impacts of 'food security' and 'hunger' – although connected, could differ, with the former being more chronic and related to availability of individuals to attend CBE (as opposed to working in agriculture), and the latter – hunger, more connected to attention, cognitive performance issues and acute health issues.

Menstruation was also cited as a prominent barrier, and again, although not surprising, highlights the need, among other things, to consider the sanitation facilities at CBE centres. The menstruation kits – given to every female beneficiary should be well received and, coupled with the extensive SRHR element of the Girls Clubs raising awareness and reducing stigma, we expect a positive impact.

Safety of beneficiaries is paramount, and, as well as tackling bullying, raising awareness and reducing stigma – especially of vulnerable individuals is extremely important to the project. Guidelines are being developed to assist solutions to ensure that girls feel safe and confident at all times – including travel to and from activities, these include initiatives such as buddying/chaperoning and travelling together where possible, involving the mother's groups and local authorities to drive a community specific and inclusive approach to this. Intensive safeguarding and stakeholder reporting mechanisms are designed to inform the approaches also.

Logframe

It is expected that targets for IOs will need to be reviewed and analysed in recognition of some of the findings, especially relating to the transition pathways and expectation/wishes of younger girls to go into vocational training and loan groups instead of continuing into formal education.

We continue to believe that the project is extremely ambitious in its scope - however this is only reflective of the complexity and scope of the issues faced by our target girls. As we learn and understand more through working with them, their families and their communities – we will further recognise and understand how the project response can adapt to meet challenges.

Annex 15: Additional tables for index items). Girls with high SRHR understanding scores had statistically higher average literacy and numeracy scores than did girls with low SRHR understanding scores. Similarly, girls with high self-confidence and life skills scores had statistically significantly higher literacy and numeracy scores than did girls with low self-confidence and life skills scores. Girls in households with high CP support scores had statistically significantly higher average literacy and numeracy scores than did girls in households with low CP support scores.

Supplemental Table 7: Learning scores by IOs

IO	Score category	Average literacy score (aggregate)	Average numeracy score (aggregate)
IO2.1: Number of girls with improved understanding of SRHR[62]	Low SRHR understanding	14.96	25.07
	High SRHR understanding*	21.53	39.51
IO2.2: Number of girls with improved self-esteem, self-confidence and well-being[63, 64]	Low self-esteem	17.18	30.03
	High self-esteem	19.12	34.17
	Low self-confidence	13.62	23.49
	High self-confidence*	19.91	35.52
IO4.2: Improved community support for SRHR and CP[65, 66]	Low SRHR support	18.62	30.93
	High SRHR support	18.28	33.16
	Low CP support	12.25	24.27
	High CP support*	21.19	36.01
IO4.3: Improved community support for girls' education through CBEs and primary school[67]	Low girls' education support	17.14	32.36
	High girls' education support	19.44	32.25
O1.3: Number of highly marginalised girls supported by GEC with improved life skills outcomes[68]	Low life skills	14.86	25.53
	High life skills*	21.89	39.62

Note: One asterisk (*) denotes differences between score categories that are statistically significant at $p < 0.05$.

A multiple linear regression model was used to examine the relative influence of barriers and marginalisation characteristics on girls' learning outcomes, namely literacy and numeracy. The results show that girls' age is the strongest predictor of performance. Older girls outperform

[62] High SRHR understanding includes girls whose score on the index was at or above 4.00 (the median of the index) on an 18.00-point scale. Low SRHR understanding includes girls whose score is below 4.00.
[63] High self-esteem includes girls whose score on the index is at or above 1.50 (the midpoint and median) on a 3.00-point scale. High self-confidence includes girls whose score on the index is at or above 1.50 (the median) of the 3.00-point scale.
[64] In conversations with TEAM Girl Malawi prior to the baseline, the team indicated that well-being would be removed form Indicator IO2.2. As a result, data for this domain was not captured at the baseline evaluation point.
[65] High SRHR support group includes households with scores at or above 2.20 (the median) on a 4.00-point scale and low SRHR support group includes households with scores below 2.20. High CP support group includes households with scores at or above 2.69 (the median) on a 3.00-point scale, and low CP support group includes households with scores below 2.69.
[66] The relationship between low and high SRHR groups may not be significant due to variability in items underlying the index. See Section 0 and Section IO4: Improvement in community members' understanding and use of support mechanisms for marginalised girls for additional details.
[67] High girls' education support group includes households with scores at or above 11.00 (the median) on a 15.00-point scale, and low girls' education support includes households with scores below 11.00.
[68] High life skills group includes girls with life skills index scores at or above 1.75 (the median) on a 3.00-point scale, and low life-skills group includes girls whose score is below 1.75 on the index.

younger girls when controlling for marital status, caregiver status, orphan-hood status, head of household status, chore burden, functional difficulty, bullying, parental support, menstruation, hunger and safety commuting to school. In the model with multiple barriers, marginalisation characteristics and demographic variables included, there are no differences in girls' performance by district. Additionally, girls who were bullied had significantly lower literacy and numeracy scores than did girls who did not report being bullied.

The model was further expanded to include girls' scores on the indices for IO2.1, IO2.2, IO4.2, IO4.3 and O1.3. From the first model—which includes barriers and marginalisation characteristics—STS retained the significant barriers, namely, bullying and age group. The influence of a girls' score on the IO is therefore examined after accounting for the effect of bullying and age group. The results show that older girls still outperform younger girls on learning outcomes, even after controlling for their level of SRHR understanding, self-esteem and confidence, SRHR support and life skills. Additionally, girls with higher community CP support (IO4.2) also had higher literacy and numeracy scores.

6.3 Transition outcome

TEAM Girl Malawi's second outcome is transition through key stages of education, training or employment. This section will present baseline findings that relate to the following indicators:

- O2.1: Number of highly marginalised girls who have transitioned into primary school
- O2.2: Number of highly marginalised girls who have transitioned into vocational training relevant to the pursuit of their career
- O2.3: Number of highly marginalised girls who have transitioned into safe, fairly paid employment or self-employment
- O2.4: Quality of life for girls who choose not to pursue vocational, business training or primary school pathways

Transition pathway age groups differ from the quantitative sample age groups articulated in the MEL framework and inception report. In this section, transition age groups are utilised for analysis purposes, while age groups throughout the remainder of the report use those groups outlined in the MEL framework and inception report (see Supplemental Table 43 for relationship between quantitative sample age groups and transition pathway groups).

Project to complete

Complete the table overleaf by outlining the transition pathways for your main intervention pathway groups.

Table 19: Transition pathways

Intervention pathway tracked for transition	Please describe the possible transition pathways for this group	Aim for girls' transition for next evaluation point	Aim for girls' transition level by the time project stops working with cohort
Transition group A (girls aged 10–15 at end of 2 years of CBE)	(Re)enrol in primary school at standard 5. Return to current situation but with essential life skills for better quality of life (to	Enrols into school. If above fails, uses life skills gained through the project to enjoy improved quality of life (to be	Enrols into school or continues to be in school and progressing through the relevant standards.

Intervention pathway tracked for transition	Please describe the possible transition pathways for this group	Aim for girls' transition for next evaluation point	Aim for girls' transition level by the time project stops working with cohort
	be defined following baseline e.g. will have acquired essential life skills to negotiate power in the household and access other protection and provision services.)	defined following baseline).	If above fails, uses life skills gained through the project to enjoy improved quality of life (to be defined following baseline).
Transition group B (girls aged 16–17 at end of 2 years of CBE)	(Re)enrol in primary school at standard 5. Enrol in vocational training. Return to current situation, but with essential life skills for better quality of life (to be defined following baseline).	Enrols into school. Enrols into vocational training, such as sewing training with Supreme Sanitary Pads; other options to be developed by project. If above fails, uses life skills gained through the project to enjoy improved quality of life (to be defined following baseline).	Enrols into school or continues to be in school and progressing through the relevant standards. Using skills gained in vocational training to access safe, fairly paid employment. If above fails, uses life skills gained through the project to enjoy improved quality of life (to be defined following baseline).
Transition group C (girls aged 18–19 at end of 2 years of CBE)	(Re)enrol in primary school at standard 5. Enrol in vocational training. Transition into safe, fairly paid employment or self-employment as part of a loan group. Return to current situation, but with essential life skills for better quality of life (to be defined following baseline).	Enrols into school. Enrols into vocational training, such as sewing training with Supreme Sanitary Pads; other options to be developed by project. Enrols in MicroLoan training and joins MicroLoan group to start own business. If above fails, uses life skills gained through the project to enjoy improved quality of life (to be defined following baseline).	Enrols into school or continues to be in school and progressing through the relevant standards. Using skills gained in vocational training and/or MicroLoans to access safe, fairly paid employment. If above fails, uses life skills gained through the project to enjoy improved quality of life (to be defined following baseline).

Pathway analysis

The transition pathways analyses at baseline are based on the age groups presented in Table 19. Because the baseline took place 2 years prior to girls' anticipated transition, the age groups used for transition pathway analyses at baseline represent the target age ranges in Table 19

subtracted by 2 years.[69] At baseline, girls were asked about their intentions to complete CBE and their hopes for themselves after CBE. Girls were asked if they believe they will finish CBE; 96.40% said yes, 1.39% said no and 2.22% said they did not know (Supplemental Table 42). For those girls that responded that they did believe they would finish CBE, they were asked what they hoped to do upon completing CBE.[70, 71] Their responses are summarised in Supplemental Table 8.

Supplemental Table 8: Girls' hopes after completing CBE (%)

Subgroup	N	Go to primary school	Go to vocational training	Work in a safe, fairly paid job	Become self-employed	Get married and care for my family	Do not know
All girls	348	20.69%	52.87%	39.08%	31.90%	14.37%	5.17%
Lilongwe	73	24.71%	57.47%	48.28%	27.01%	17.82%	2.87%
Dedza	174	31.51%	63.01%	43.84%	30.14%	24.66%	1.37%
Mchinji	101	5.94%	37.62%	19.80%	41.58%	0.99%	11.88%
Age 10–11	54	29.63%	38.89%	35.19%	25.93%	20.37%	12.96%
Age 12–16	197	21.32%	53.81%	38.58%	35.03%	12.69%	4.57%
Age 17–19	97	14.43%	58.76%	42.27%	28.87%	14.43%	2.06%
Transition pathway							
Transition group A (girls aged 10–15 at end of 2 years of CBE)	96	30.21%	40.63%	36.46%	27.08%	17.71%	9.38%
Transition group B (girls aged 16–17 at end of 2 years of CBE)	105	21.90%	55.24%	40.95%	40.95%	12.38%	3.81%
Transition group C (girls aged 18–19 at end of 2 years of CBE)	147	13.61%	59.18%	39.46%	28.57%	13.61%	3.40%
Subgroup							
Married and caregiver	70	10.00%	50.00%	40.00%	34.29%	12.86%	4.29%
Orphaned	95	22.11%	57.89%	40.00%	30.53%	18.95%	4.21%
Head of household	10	0.00%	50.00%	60.00%	10.00%	0.00%	0.00%
High chore burden	168	14.88%	50.00%	39.29%	24.40%	8.93%	4.76%

[69] Transition pathway group 1 includes girls aged 10-13 at baseline, transition pathway group 2 includes girls aged 14-15 at baseline and transition pathway group 3 includes girls older than 15 at baseline.
[70] Percentages across characteristics do not equal 100.00%. Girls could provide multiple responses to the question. All response options, except for 'other', are included in Supplemental Table 8.
[71] Vocational training was explained to respondents as per TEAM Girl Malawi's definition, based on what the training will include through the project and/or what is generally understood as vocational training in Malawi.

Subgroup	N	Go to primary school	Go to vocational training	Work in a safe, fairly paid job	Become self-employed	Get married and care for my family	Do not know
Girls with disabilities	118	22.88%	49.15%	37.29%	38.14%	16.10%	5.93%
Barrier							
Bullying	24	25.00%	41.67%	37.50%	16.67%	12.50%	8.33%
School cost	318	20.75%	52.83%	38.05%	32.70%	14.47%	5.66%
Parent support	65	15.38%	49.23%	35.38%	41.54%	18.46%	7.69%
Menstruation	106	24.53%	50.00%	46.23%	39.62%	18.87%	4.72%
Food insecure or hunger	224	19.64%	51.79%	35.71%	33.48%	14.73%	4.91%
School safety	96	13.54%	51.04%	36.46%	36.46%	11.46%	5.21%

Of the 348 girls who believed they will complete CBE, more than half (52.87%) reported that they hoped to go to vocational training, 39.08% said they wanted to work in a safe, fairly paid job and just 20.69% hoped to (re)enrol in primary school.[72] About one in 5 responded that they hoped to go to primary school following completion of CBE. Overall, larger proportions of girls in younger age groups reported that they hoped to return to primary school after competing CBE. Larger proportions of girls in older age groups reported that they hope to go to vocational training or work in safe, fairly paid jobs. These trends align with TEAM Girl Malawi transition. While these proportions are girls' intentions for transition, at subsequent timepoints both their intentions and actual transition rates collected through monitoring data will be reported.

Headline analysis

Girls' educational status at baseline—populated using the TEAM Girl Malawi enrolment database—is presented in Table 20. A majority of girls (84.55%) reported attending school in the past but dropping out. The same proportion of girls (28.49%) reached standard 3 and as did standard 4, while 14.24% of girls had never been to school prior to enrolling in TEAM Girl Malawi. STS considered a girl 'currently employed' if she responded that she completed a paid activity on a typical day. This categorisation applied to 5.18% of girls. None of the girls were enrolled in formal school when they were enrolled in TEAM Girl Malawi.

Table 20: Status at baseline (intervention population)

Status	Intervention (%)
Never been to school[73]	14.24%
Been to school but dropped out[74]	84.55%
Standard reached—standard 1	11.01%
Standard reached—standard 2	17.78%
Standard reached—standard 3	28.49%
Standard reached—standard 4	28.49%

[72] When asked on the household survey, 49.14% of parents said they hoped their girls transitioned into vocational training, and 48.84% said they would like their girl to achieve an upper secondary-level education.
[73] Intervention proportion calculated using TEAM Girl Malawi enrolment database for girls who did not complete at least standard 1.
[74] Calculated using TEAM Girl Malawi enrolment database for all girls who completed at least standard 1.

Currently enrolled in formal school	0.00%
Currently employed[75]	5.18%
Source: Intervention N = 2008	TEAM Girl Malawi enrolment database

Characteristic subgroups and barrier analysis

STS analysed the status of sampled girls at baseline by transition pathway, subgroup and barrier. Results are presented in Supplemental Table 9. Overall, the largest proportion of girls reached standard 3—31.75%; 29.63% reached standard 4. About one in 10 girls from the baseline sample had never been to school prior to enrolling in TEAM Girl Malawi. A smaller proportion of girls in Mchinji had never been to school, and a larger proportion of girls in Mchinji reported being currently employed at the start of TEAM Girl Malawi than in other districts or overall. Larger proportions of girls from young age groups (10–11 and 10–15) reported having never been to school. Notably, 37.60% of girls who reported experiencing food insecurity or hunger had never been to school.[76] Also notable was that 41.33% of girls with disabilities reported having reached standard 4.

Supplemental Table 9: Status at baseline by subgroups (sample)[77]

Subgroup	N	Never been to school	Standard 1	Standard 2	Standard 3	Standard 4	Currently employed
All girls	**378**	**10.32%**	**9.26%**	**19.05%**	**31.75%**	**29.63%**	**0.79%**
Lilongwe	77	10.39%	12.99%	32.47%	31.17%	12.99%	0.00%
Dedza	193	11.40%	22.28%	33.16%	26.94%	6.22%	0.52%
Mchinji	108	4.63%	17.59%	28.70%	33.33%	15.74%	2.60%
Age 10-11	60	25.00%	31.67%	25.00%	16.67%	1.67%	0.00%
Age 12-16	216	8.33%	6.02%	19.44%	32.87%	33.33%	0.93%
Age 17-19	102	5.88%	2.94%	14.71%	38.24%	38.24%	0.98%
Transition pathway							
Transition group A (girls aged 10–15 at end of 2 years of CBE)	106	12.00%	28.41%	30.68%	26.14%	14.77%	1.00%
Transition group B (girls aged 16–17 at end of 2 years of CBE)	113	6.36%	3.88%	22.33%	37.86%	35.92%	0.00%
Transition group C (girls aged 18–19 at end of 2 years of CBE)	159	1.99%	4.05%	14.86%	39.19%	41.89%	1.32%

[75] Intervention proportion calculated using TEAM Girl Malawi enrolment database for girls who reported completing a paid activity on a typical day.
[76] Hunger is used as a proxy measure for girls who experience extreme poverty within a high poverty population.
[77] Percentages across characteristics do not equal 100.00%. Girls schooling status and employment status were asked as two separate questions. As a result, a girl may be counted in schooling status columns and employment status columns.

Subgroup	N	Never been to school	Standard 1	Standard 2	Standard 3	Standard 4	Currently employed
Subgroup							
Married and caregiver	75	9.33%	1.33%	10.67%	37.33%	41.33%	0.00%
Orphaned	99	5.05%	5.05%	24.24%	31.31%	34.34%	1.01%
Head of household	11	0.00%	0.00%	9.09%	36.36%	54.55%	0.00%
Head of household	181	11.05%	4.97%	20.99%	31.49%	31.49%	1.66%
High chore burden	130	11.54%	6.92%	14.62%	31.54%	35.38%	0.00%
Girls with disabilities	75	9.33%	1.33%	10.67%	37.33%	41.33%	0.00%
Barrier							
Bullying	29	13.79%	10.34%	20.69%	34.48%	20.69%	0.00%
School cost	334	9.58%	9.58%	19.16%	32.34%	29.34%	0.90%
Parent support	70	10.00%	11.43%	15.71%	35.71%	27.14%	0.00%
Menstruation	125	10.40%	5.60%	12.00%	34.40%	37.60%	0.00%
Food insecure or hunger	236	37.60%	10.17%	17.80%	33.47%	27.97%	0.85%
School safety	103	12.62%	8.74%	12.62%	29.13%	36.89%	0.00%

Indicator O2.4 was included in the TEAM Girl Malawi logframe to ensure that even if girls do not transition into one of the pathways outlined by the project their quality of life is measured and improved to ensure they are better off. At baseline, qualitative data was analysed to determine what domains of quality of life would be valued by beneficiaries and stakeholders. Based on baseline findings, the following topics will be explored through quantitative and qualitative tools in future evaluation points to enable measurement of girls' quality of life as defined by them:

- Access to livelihoods or income generating activities
- Access to financial resources and capital
- Strength of social networks and support systems
- Marital status
- Health status and access to health services generally and for SRHR specifically
- Level of engagement in risky behaviours, such as sex work, alcohol and drug consumption or gambling
- Self-esteem, self-confidence and self-determination as well as social emotional skills
- Safety and freedom from violence, including bullying, harassment, physical violence or sexual violence

Transition pathways analyses for learning outcomes and intermediate outcomes

Learning scores by transition groups are presented in Supplemental Table 10. Overall, average aggregate literacy and numeracy scores were statistically higher than transition group A for transition group B and transition group C. Given that ages are determinants of transition pathway

groups, these findings are similar to those presented in Table 18 for the quantitative sample age groups, in which older age groups outperformed the youngest age group.

Supplemental Table 10: Learning scores by transition pathway group

	Average literacy score (aggregate)	Average numeracy score (aggregate)
All girls	**18.31**	**32.23**
Transition group A	12.27	17.95
Transition group B^	19.61	37.18
Transition group C^	21.41	38.24

Note: One caret (^) denotes that the difference between the group and transition group A is statistically significant at $p < 0.05$.

Baseline findings for IO2.1 and IO2.2 are presented in Supplemental Table 11 and Supplemental Table 12, respectively.[78] Girls in transition group A had the greatest proportion of low scores (78.00%) on the SRHR understanding index, while transition group C had the greatest proportion of high scores (75.50%) on the SRHR understanding index. There were no differences between transition groups and the quantitative sample age groups. For IO2.2 self-esteem and self-confidence results, there were no differences between age groups. By comparison, using the sampling age groups, girls in the oldest age group had the greatest proportion of girls in the high score category.

Supplemental Table 11: IO2.1 SRHR understanding results by transition pathway group[79]

Category	N	Score	Proportion of total
All girls	**361**	**Low score**	**48.20%**
		High score	**51.80%**
Transition group A	100	Low score	78.00%
		High score	22.00%
Transition group B	110	Low score	53.64%
		High score	46.36%
Transition group C	151	Low score	24.50%
		High score	75.50%

Supplemental Table 12: IO2.2 Self-esteem and self-confidence results by transition pathway group[80]

Category	N	Score	Proportion of total (Self-esteem)	Proportion of total (Self-confidence)
All girls	**361**	**Low score**	**39.06%**	**24.65%**
		High score	**60.94%**	**75.35%**
Transition group A	100	Low score	39.00%	37.00%
		High score	61.00%	63.00%
Transition group B	110	Low score	46.36%	21.82%
		High score	53.64%	78.18%
Transition group C	151	Low score	33.77%	18.54%

[78] See Section IO2: Sexual and reproductive health and rights, self-confidence, self-esteem and wellbeing for an in-depth description and analysis of IO2.
[79] For IO2.1 SRHR understanding, the cut-off point for low and high scores is the median score, 4.00. Girls who scored 4.00 or higher were categorised as 'high scores'. Girls who scored less than the median of 4.00 were categorised as 'low scores'.
[80] High self-esteem scores were defined as scores above 1.50, the median of the self-esteem index. High self-confidence scores were defined as scores above 1.50, the median of the self-confidence index.

Category	N	Score	Proportion of total (Self-esteem)	Proportion of total (Self-confidence)
		High score	66.23%	81.46%

Baseline findings for IO4.2 and IO4.3 are presented in Supplemental Table 13, Supplemental Table 14 and Supplemental Table 15.[81] Perceptions of community support for SRHR were comparable across the three transition groups and perceptions of CP support were statistically significantly higher for transition group A and B than for transition group C.

Supplemental Table 13: IO4.2 SRHR support mean scores by transition pathway group

Category	Disaggregation	N/n	Mean score (on 4-pt scale)
All households	**NA**	**348**	**2.18**
Transition group	Transition group A	102	2.13
	Transition group B	106	2.18
	Transition group C	140	2.22

Supplemental Table 14: IO4.2 Child protection support mean scores by transition pathway group

Category	Disaggregation	N/n	Mean score (on 3-pt scale)
All households	**NA**	**371**	**2.60**
Transition group	Transition group A	104	2.38
	Transition group B^	112	2.63
	Transition group C^	155	2.74

Note: One caret (^) denotes that the difference between the group and transition group A is statistically significant at $p < 0.05$.

Supplemental Table 15: IO4.3 Girls' education support mean scores by transition pathway group

Category	Disaggregation	N/n	Mean score (out of 15.00)
All households	**NA**	**370**	**10.05**
Transition group	Transition group A	104	10.57
	Transition group B	111	10.14
	Transition group C	155	9.65

Supplemental Table 16 presents findings for O1.3.[82] Transition group C had the greatest proportion of girls with high life skills scores.

Supplemental Table 16: O1.3 Results by transition pathway group[83]

Category	N	Score	Proportion of total
All girls	**361**	**Low score**	**50.14%**
		High score	**49.86%**
Transition group A	100	Low score	72.00%
		High score	28.00%
Transition group B	110	Low score	53.64%

[81] See Section IO4: Improvement in community members' understanding and use of support mechanisms for marginalised girls for an in-depth description and analysis of IO4.
[82] See Section 7.2 for an in-depth description and analysis of O1.3.
[83] High life skills scores were defined as scores greater than 1.74, the median of the life skills index.

Category	N	Score	Proportion of total
		High score	46.36%
Transition group C	151	Low score	33.11%
		High score	66.89%

6.4 Sustainability outcome

Baseline evidence on O3 Sustainability is presented in the following section for system, community and learning space indicators and primarily draws upon qualitative data.

System

KIIs were conducted with 8 government officials—6 at the district level and 2 at the national level—to understand conditions for sustainability at baseline. Overall, government officials were generally knowledgeable about plans and policies, able to name the formal name, informal name or parts of the plan or policy. Results suggest, however, that levels of knowledge vary and that individuals are differently informed based on their position—technical versus support staff—and level—national versus district. One national MOEST official was highly informed of policies, while a different MOEST official admitted to not being conversant in the National Girls' Education Strategy nor the National Girls' Education Communication Strategy. One official said that the government supports the dissemination of policies through community sensitisation, which is carried out by primary education advisers and head teachers. He also said that dissemination of policies is tracked by having officials interview individuals at schools about their awareness and implementation of policies.

Government officials identified concerns over a lack of resources to implement policies that support marginalised girls' education, including understaffed offices, logistical challenges, low technical capacity of staff, lack of technology and limited resources to monitor dissemination and application. Officials were, however, able to name several different mechanisms to improve support, suggesting that they have ideas for how to motivate their colleagues to better engage with marginalised girls' education. Examples include presentations, CBE site visits, trainings, capacity-building activities and sensitisation campaigns.

District education officials mentioned examples of how their offices respond to the needs of marginalised girls. Namely, officials mentioned that there are policies in place to support marginalised girls and that those plans are disseminated to schools and communities. They also described the role of the special needs education district coordinator, who trains teachers on inclusive education. Finally, they mentioned that district education offices encourage schools to have reporting systems—such as a CP committee—where learners can report ethical issues and proper actions can be taken.

Given this evidence, the proposed system-sustainability score at baseline using the sustainability scorecard is 1.00 out of 4.00. There is evidence of foundational support for sustainability, but there does not appear to be consistent conditions in place across government offices and levels. It is also unclear to what extent the actions that support marginalised girls' education have been implemented or how effective dissemination efforts of policies and plans have been.

Community

At baseline, parents and caregivers were asked about their participation in school improvement meetings at their local primary school. Findings indicate that 10.92% had participated. Of those

that participated, 28.95% suggested prioritising support for marginalised girls' learning, and 42.11% suggested prioritising special learning resources or supports. Only 5.26% suggested prioritising access for girls with disabilities, and 21.05% said they did not suggest any improvement priorities (Supplemental Table 14).

Parents, caregivers, community leaders and mothers' group members expressed strong support for marginalised girls' education through schools, vocational training and business training. When asked on the household survey, 49.14% of parents and caregivers said they hoped their girls transitioned into vocational training, and 48.84% said they would like their girl to achieve an upper secondary-level education. Further, when asked if they agreed that even when funds are limited, it is worth investing in their girl's education, 87.07% of parents and caregivers strongly agreed (Supplemental Table 40).

Further, key community leaders and mothers' group members were able to provide several examples of ways they encourage members of their community to actively support marginalised girls. These include:

- Lead by example: Encouraging community members by demonstrating support for marginalised girls
- Communication to parents and caregivers: Having meetings with parents, caregivers, teachers and school committees to build a solid chain of communication between stakeholders in girls' education
- Community meetings: Inviting parents, caregivers and girls to community meetings to share information about NGO activities like TEAM Girl Malawi
- Whole-community approaches: Engaging the entire community with development activities and girls' education initiatives to create inclusive environment; considering those who do not participate as 'enemies' of development
- Home visits: Conducting visits to households with girls who are not attending school due to disabilities or pregnancy
- Material support: Soliciting and distributing community contributions so that households with high levels of poverty can buy school uniforms and supplies
- Penalties and fines: Enforcing laws on early marriage and fining households that do not allow girls to attend school

Community leaders and mothers' group members appeared to understand CP reporting systems and sources of support available to survivors of abuse. Respondents shared that in cases of child rape community police, village chiefs and heads and health services were engaged. In one FGD, a participant described the presence of a CP committee in the community. Community leaders also described the role of CP committees in KIIs. Mothers' group members described the role of school management committees and *ikata*, or community policing groups, in cases of school-based violence. In cases of child abuse perpetrated by caregivers, one mothers' group member indicated that, when she witnesses physical violence such as slapping or hitting, she informs children to report their caregivers.

Given the evidence, the proposed community sustainability score at baseline is 2.00 out of 4. There is evidence of strong foundational support for girls' education and protection. Efforts should be made to ensure that community reporting mechanisms are strengthened and that community members are utilising mechanisms consistently and effectively.

Learning space

At baseline, no evidence was collected from primary schools or CBEs as girls had not yet begun formal learning sessions. As a result, findings for the learning space sustainability indicators are limited.

In KIIs, government officials referenced school improvement grants that support orphaned and vulnerable children. Additionally, they mentioned that early grade teachers receive training on inclusive education techniques and are supported by specialists. They also shared that learners are screened for disabilities and provided with assistive devices such as hearing aids. However, it is unclear to what extent primary schools are using these supports.

School safety, including on the way to school and at school, was an expressed barrier by girls in FGDs and surveys. Specifically, on the girls survey, respondents were asked if they felt safe travelling to and from school; 46.15% reported that they did not. Additionally, 23.08% of girls reported feeling unsafe at school (Supplemental Table 35). Of the 353 caregivers responding to the household survey, 24.08% reported that it was 'fairly unsafe' or 'very unsafe' for girls to travel to schools in their area. Of the 316 caregivers who reported that their girl was not enrolled in primary school, 11.71% reported that she was not enrolled because it is unsafe to travel to or from school in that area; 8.54% because their girl is not safe at school (Supplemental Table 40. Girls also described community-level barriers to attendance that were primarily related to safety concerns, including accidents with vehicles or motorbikes and fear of violence including gender-based violence. Girls also expressed concern about safety in the learning environment. They mentioned fears of bullying, fighting and physical violence at schools as barriers to a quality learning environment.

Evidence on inclusive teaching practices were limited because learning sessions had not yet started. In FGDs, girls did not expressly discuss or provide examples of inclusive teaching practices or differentiated treatment—either positive or negative—in the classroom.

Given the evidence, the proposed learning space sustainability score at baseline is 0.00 out of 4.00. There was not enough evidence to conclude that conditions for sustainability in the learning space exist at baseline.

Table 21: Sustainability indicators

	System	Community	Learning space
Indicator 1:	Education officials are informed about the policies that can support marginalised girls, including monitoring and measuring support mechanisms for girls at the district and national level. **Results:** Government officials were generally knowledgeable about plans and policies, though levels of knowledge vary based on positions.	A wide cross-section of the community participates in school improvement planning. **Results:** 10.92% of parents and caregivers had participated in school improvement meetings at their local primary school.	Primary schools allocate resources (time, staff, funding) to enable marginalised adolescent girls to attend and learn. **Results:** NA
Indicator 2:	Education officials can explain how they might motivate people to be more	A wide cross-section of the community advocates for	Girls report that schools are safe.

	System	Community	Learning space
	active supporters of marginalised girls. **Results:** Officials could name several different mechanisms to improve support.	improvement plan targets and budgets which meet the needs of marginalised adolescent girls. **Results:** 21.05% of parents and caregivers that attended meetings did not suggest any improvement priorities.	**Results:** 46.15% of girls reported that they did not, in the past, feel safe travelling to and from school; 23.08% of girls reported feeling unsafe at school in the past.
Indicator 3:	District education offices respond to the needs of marginalised girls. **Results:** District education officials could mention examples of how their offices respond to the needs of marginalised girls.	Key community leaders and a critical mass of stakeholders are supportive of marginalised girls attending learning centres, vocational training, or business training. **Results:** Parents, caregivers, community leaders and mothers' group members expressed strong support for marginalised girls' education through schools, vocational training and business training.	Girls report that teachers and teaching is inclusive. **Results:** NA
Indicator 4:		Key community leaders and a critical mass of stakeholders can explain how they would go about trying to motivate people in their community to more actively support marginalised girls. **Results:** Key community leaders and mothers' group members were able to provide several examples of ways they encourage members of their community to actively support marginalised girls.	

	System	Community	Learning space
Indicator 5:		Key community leaders and mothers' group members understand the CP reporting systems and sources of support available to survivors of abuse. **Results:** Community leaders and mothers' group members appeared to understand CP reporting systems and sources of support available to survivors of abuse.	
Baseline Sustainability Score (0–4)	**1.00**	**2.00**	**0.00**
Overall Sustainability Score (average of the level scores)	**1.00**		

Project to complete

Complete the table below by answering the questions in the table. Once completed, provide narrative analysis of the points raised in the table to explain the change the project intends to achieve. Ensure your analysis reflects the scores your external evaluator rated for each of your sustainability indicators.

Table 22: Changes needed for sustainability

Questions to answer	System	Community	Learning space	Family or household	Girl
Change: What change should happen by the end of the implementation period	Improved capacity of local officials to support girls' education through existing functions and adopting new approaches**.** Police and statutory agencies are delivering their roles in CP reporting with support from project staff. Local or national government has engaged and understood evidence from the project through quarterly National and District Steering Committee meetings and field visits. Social enterprises and private sector actors, like Supreme Sanitary Pads and Microloan Foundation, engage to improve girls' opportunities for skills and transition to	Key community leaders and a critical mass (NB the FM has asked us to clarify 'critical mass') of stakeholders are convinced of the benefits and are supportive of marginalised girls attending learning centres, vocational training or business training. Key community leaders, including mothers' groups and traditional authorities, have the capacity to independently run community listening clubs and make CP reporting systems run effectively A critical mass of stakeholders representing all sections of the community actively participate in school improvement planning that addresses the	Head teacher and a critical mass of teachers and staff in CBEs and schools are convinced of the benefits and have the capacity to deliver changed practice independently. To the extent possible, existing financial and other resources, including school improvement grants and microloans, are being used.	Husbands, parents and caregivers are convinced of the benefits and are supportive of marginalised girls attending learning centres, vocational training or business training.	Girls are safe and have increased literacy and numeracy levels and increased SRHR, self-esteem and confidence skills to earn and access to loans.

Questions to answer	System	Community	Learning space	Family or household	Girl
	employment, including self-employment	needs of marginalised girls. Financial and other resources are increasingly being mobilised locally through parents' and caregivers' participation in loan groups or through school improvement grants.			
Activities: What activities are aimed at this change?	Training government staff on school review, support visits to schools and CBEs and teacher training. Child protection training for statutory agencies. Steering Committees. MLF loan groups and Supreme vocational training piloted and established in target communities.	Community sensitisation and participation in adaptive management meetings. Training mothers' groups and learning centre management committees on education rights. Training school leaders and modelling inclusive school review and school improvement planning. Caregivers' participation in loan groups and vocational training.	Leadership training for headteachers and district education officers on the National Education Standards and School Review process. School review and community meetings to assess schools against the national education standards. Primary teachers trained on gender-responsive and inclusive education.	Community sensitisation and participation in adaptive management meetings, interaction with trained mother group and LCMCs, participation in discussions at School Review and School Improvement Planning community meetings. Skills gained at vocational training. Access to microloans and membership of loan groups.	CBEs, girls' clubs, vocational training, microloan groups, school review and school improvement planning, capacity building of CP systems.

Questions to answer	System	Community	Learning space	Family or household	Girl
Stakeholders: Who are the relevant stakeholders?	MOEST and MOGCDSW MOEST's CBE and special needs department MOEST Primary Education Advisers at the district level MOGCDSW and district-based CP officers Social welfare officers Local police	Community leaders Mothers' group members LCMC members Traditional authorities School governing bodies Alumnae girls	School leaders Primary education advisers Teachers	Husbands Parents and caregivers Community leaders Mothers' group members LCMC members Traditional authorities School governing bodies	Beneficiary boys Families of beneficiary girls CBE facilitators Learning assistants Agents of change Vocational trainers Loan group leaders Teachers and headteachers Alumnae Girls LCMCs Mothers' groups Statutory agencies Police Community leaders
Factors: What factors are hindering or helping achieve changes? Think of people, systems, social norms, etc.	Budget constraints: It is unlikely that additional government funding will be available to scale up or sustain CBEs. Instead the project aims to build capacity to make mainstream	Time for stakeholders to participate and engage others. Social norms that are unsupportive of marginalised girls' learning may not fully	Inclusive education requires leadership from head teachers and support from primary education advisers. Teachers need sufficient skills and	Time to participate. Social norms that are unsupportive of marginalised girls' learning may not fully change during the project lifetime.	Effectiveness of facilitator, learning assistant, AoC, vocational trainer and loan group leader training. Effectiveness of CBE, girls' club, vocational

Questions to answer	System	Community	Learning space	Family or household	Girl
	schools inclusive. At the system level this relies on engagement of MOEST officials who will support schools and MOGCDSW officials who will support communities to be safe and enabling environments for marginalised girls' learning. Government budget constraints may impact on these agencies. Capacity to conduct school review and follow up on CP cases: Especially regarding CP, project staffing and resources still play role but there is potential for this be phased out.	change during the project lifetime.	confidence to share these skills with peers. Community support is required for the allocation of school funds to support inclusive education. Ongoing staff capacity development requires budget and time allocation by district government for primary education advisers.	Effectiveness of vocational training and loans on family poverty reduction.	training and loan group curriculum. Effectiveness of CP systems. Impact of community sensitisation on social norms**.**

7. Key intermediate outcome findings

Baseline results for the following TEAM Girl Malawi IOs are presented in this section:

- IO1: Attendance
- IO2: Sexual and reproductive health and rights, self-confidence, self-esteem and wellbeing
- IO3Improvement in quality of education at CBE centres and primary schools
- IO4: Improvement in community members' understanding and use of support mechanisms for marginalised girls
- IO5: Strengthened district and national leadership and engagement in marginalised adolescent girls' education

Additionally, key findings on indicator O1.3 Number of highly marginalised girls supported by GEC with improved life skills outcomes are presented in Section 7.2.

7.1 Key intermediate outcome findings

IO1: Attendance

TEAM Girl Malawi's first IO is attendance. Specifically, the project's ToC assumes that improved attendance to sites of learning is a prerequisite for better learning, transition and sustainability for marginalised girls.

IO1 indicators and relevant baseline information are detailed in Table 23. Because learning had not yet begun at project CBEs at the time of data collection, all IO1 indicators are set at zero for baseline. Qualitative findings under this IO will provide critical feedback to the project about how to support attendance over the years of the project.

Table 23: IO1 Attendance indicators

IO	IO indicator	Sampling and measuring technique used	Who collected the data?	Baseline level	Target for next evaluation point	Will IO indicator be used for next evaluation point? (Y/N)
IO1: Attendance	IO1.1: Percentage of beneficiaries, teachers and educators and caregivers who have positive views on how the support received has helped reduce the barriers to regular attendance	Girls survey Household survey CBE facilitator survey	NA at baseline	0%	60%	Yes

IO	IO indicator	Sampling and measuring technique used	Who collected the data?	Baseline level	Target for next evaluation point	Will IO indicator be used for next evaluation point? (Y/N)
IO1: Attendance	IO1.2: Number of project girls and boys with identified marginalisation characteristics (as defined by the project) who maintain 80% attendance records CBEs or Girls' Clubs	Project attendance Classroom observation	NA at baseline	Girls – 0	Girls – 2656 (80% of beneficiaries)	Yes
				Boys – 0	Boys – 531 (80% of beneficiaries)	
IO1: Attendance	IO1.3: Number of project girls and boys with identified marginalisation characteristics (as defined by the project) regularly attending vocational and business training initiatives	Project attendance records Classroom observation	NA at baseline	Girls – 0	Girls – TBC	Yes
				Boys – 0	Boys – TBC	

Main qualitative findings

- Overall, FGDs with adolescent girls and boys identified several individual-level barriers to attendance, including sickness, lack of interest in school, dirty clothes, lack of learning materials and menstruation.
- Household-level barriers to attendance included heavy chore-burdens, poverty, hunger and lack of parental support.
- Community-level barriers to attendance were primarily related to safety concerns, including accidents with vehicles or motorbikes or fear of violence and gender-based violence.

Main findings

The qualitative data collected at baseline provided substantial insight into the perceived barriers to attendance in the TEAM Girl Malawi project that girls and boys anticipate. Because these beneficiaries have not yet started learning sessions at the CBEs, the barriers identified are likely informed by their previous experiences with the formal school system. Barriers are grouped by individual-level, household-level and community-level/outside the household. They include, but are not limited to, the barriers described in Table 15.

Barriers to attendance identified during FGDs by specific characteristic subgroups are described in Supplemental Table 17. A frequently mentioned individual-level barriers to attendance was sickness, which included getting sick or having seizures on the way to school. This potential barrier was mentioned by younger and older girls, married girls, young mothers, girl with disabilities and older boys. Lack of interest in attending school was cited across FGDs with adolescent girls, adolescent boys and mothers' groups. More specifically, adolescents appeared uninterested in attending school regardless of if their parents or caregivers encouraged them. Several groups of adolescents mentioned that dirty clothes or lack of appropriate clothes would keep them from attending. Younger and older adolescent girls said that lack of learning materials, such as exercise books or writing materials, would be a barrier to attendance. Girls also mentioned that some girls in their communities do not attend school because they were engaged in sex work or transactional sex activities. Specifically, girls said that their peers around age 15 and 16 drop out of school because they engage in prostitution.

Household-level barriers to attendance primarily related to chore-burden placed on adolescents from their parents or caregivers. Across subgroups, respondents cited chores—including instructions from parents or caregivers to do chores, care for siblings or do farm work—as a household-level barrier to attendance. Poverty was also cited frequently. Respondents said they could not or would not go school because they could not afford school fees, uniforms, exam fees or school materials.

Barriers in attendance at the community-level or outside of the household primarily concerned issues of safety and security. Across subgroups, there were concerns about being hit by vehicles, motorbikes or bicycles as well as general accidents on the way to school. Being attacked by strangers was also cited across adolescent FGDs as a potential barrier to attendance.

Supplemental Table 17: Barriers to attendance by subgroup from qualitative data

Barrier type	Barrier	Subgroup
Individual	Menstruation	Adolescent girls aged 15–19 Caregivers
Individual	Tardiness or non-attendance due to domestic chore responsibilities	Adolescent girls aged 15–19 Girls with disabilities
Individual	Hunger or food insecurity	Orphaned Adolescent boys aged 15–19
Individual	No perceived purpose or benefits of schooling	Caregivers
Individual	Being in the same classroom as current/former sexual partners	Adolescent boys aged 15–19
Individual	Prostitution or sex work	Adolescent girls aged 15–19
Individual	Violence or punishment for misbehaviour	Adolescent girls aged 15–19 Adolescent boys aged 15–19
Individual	Using drugs, smoking Indian hemp, drinking alcohol and gambling	Adolescent boys aged 15–19
Household	Relatives discouraging their attendance	Adolescent girls aged 10–14 Married Orphaned girls
Household	Husbands restricting their attendance	Married Orphaned
Household	Family illness	Adolescent boys aged 15–19 Married
Household	Being displaced from home or homeless	Orphaned Girls with disabilities

Barrier type	Barrier	Subgroup
Community	Peer pressure or discouragement from attending by friends	Adolescent girls aged 15–19 Adolescent boys aged 15–19 Orphaned
Community	Attacks or violence perpetrated by friends, bullying	Adolescent girls aged 10–14 Adolescent girls aged 15–19
Community	Fear of rape	Caregivers
Community	Snakes	Adolescent boys aged 15–19
Community	Distance from household to school	Adolescent boys aged 15–19

Feedback from community members and mothers' group members also provided informative perspectives on potential barriers to attendance. A community leader cited age as a predictor of attendance. Specifically, he was said that adolescents aged 15 through 19 are more vulnerable to be dropouts because they are more likely to explore sexual relationships, drink beer and experience peer pressures to earn a living. *Gule Wamkulu*, a cultural activity of the Chewa, was also cited as a potential interference to attendance. Mothers' group members said that boys often skip school to see video shows during the day, and one community leader said that girls do the same.

Reflections

Because learning had not yet started at the time of data collection, data collection for IO1 focused primarily on identifying potential obstacles or barriers to access to, attendance at and completion of CBE through qualitative methods. The findings under IO1 at baseline should be used to provide formative feedback to the project in order to reduce or eliminate obstacles and barriers that girls and boys may confront in the coming years of the project.

The project may want to simplify indicator IO1.1—percentage of beneficiaries, teachers, educators, and caregivers who have positive views on how the support received has helped reduce the barriers to regular attendance. Specifically, the project should assess if it is most important to measure respondents' perceptions about barriers or if it is more relevant to measure the proportion of girls experiencing barriers. An alternative indicator could be 'percentage of beneficiaries, teachers, educators and caregivers who report that barriers to regular attendance have been reduced as a result of support received'.

Specifically, STS recommends the following updates to indicators under IO1:

- IO1.2: Average attendance rate of girls and boys with identified marginalisation characteristics at CBEs/Girls' clubs
- IO1.3: Average attendance rate of and boys with identified marginalisation characteristics (as defined by the project) at vocational and business training programmes

At baseline, perceptions of barriers to attendance from qualitative data were used to report against this IO. At the next evaluation point, several new tools or items will be introduced to track indicators under IO1 using qualitative and quantitative data. The quantitative data to be added include:

For IO1.1:

- CBE facilitator surveys
- Additional girls survey items
- Additional household survey items

For IO1.2:

- Classroom observations, including headcount attendance at evaluation points and project monitoring data

Qualitative data for all IO2 indicators will be captured from adolescent girls and boys, mothers' group members, caregivers and CBE facilitators. Between evaluation periods, TEAM Girl Malawi should ensure that attendance records from CBEs and Girls' Clubs are consistently tracked and collated. The monitoring data on attendance will be combined with data collected at the next evaluation point to report on trends in attendance across the sample during the intervening months.

Targets

STS proposes the following targets for year 3:

- IO1.1 % beneficiaries', teachers/educators' and care givers' who have positive views on how the support received has helped reduce the barriers to regular attendance: 50%
- IO1.2 Average attendance rate of girls and boys with identified marginalisation characteristics at CBEs/Girls' clubs: 60%
- IO1.3 Average attendance rate of and boys with identified marginalisation characteristics (as defined by the project) at vocational and business training programmes: 60%

Project Checks on Intermediate Outcomes

Ensure that the IO analysis reflects the links between different levels in the logframe and informs the validity of the Theory of Change. This includes checking whether the EE (?) have:

- Measured and analysed all IO indicators presented in logframe.
- Disaggregated the data according to the logframe.
- Used both the qualitative and quantitative analysis stated in the logframe.
- Related the IO analysis to the analysis of Outcomes.

Project confirmation that all the IO analysis has been covered.

IO2: Sexual and reproductive health and rights, self-confidence, self-esteem and wellbeing[84]

TEAM Girl Malawi's second IO is SRHR, self-confidence, self-esteem and wellbeing. Specifically, the project's ToC assumes that improved knowledge and understanding of SRHR as well as improved self-confidence, self-esteem and well-being are prerequisites for better learning, transition and sustainability outcomes for marginalised girls.

IO2 indicators and relevant baseline information are detailed in Table 24. Because all indicators under IO2 require results to be reported as a number of girls with improved scores over baseline, only 2 results are presented at baseline: the mean score on an index and the proportion of girls categorised as having high scores defined as at or above the median score on an index.[85] At

[84] In conversations with TEAM Girl Malawi prior to the baseline, the team indicated that well-being would be removed from indicator IO2.2. As a result, data for this domain was not captured at the baseline evaluation point.
[85] High score categories at baseline are used for relative comparability of scores and should not be assumed to indicate high knowledge or performance. Comparisons between high and low score categories can, however, be used to understand performance relative to the median across subgroups.

midline, the number of girls with improved mean scores over baseline will be reported, as will the proportion of girls in the high score category.

Table 24: IO2 Sexual and reproductive health and rights, self-confidence, self-esteem and wellbeing indicators

IO	IO indicator	Sampling and measuring technique used	Who collected the data?	Baseline level	Target for next evaluation point	Will IO indicator be used for next evaluation point? (Y/N)
IO2: Sexual and reproductive health and rights, self-confidence, self-esteem and wellbeing	IO2.1: Number of girls with improved understanding of SRHR	Girls survey	STS	4.06 out of 18.00 mean score 51.80% of girls categorised as having a high score[86]	3320 83% of 4000 - cohorts 1 and 2	Yes
IO2: Sexual and reproductive health and rights, self-confidence, self-esteem and wellbeing	IO2.2: Number of girls with improved self-esteem, self-confidence and well-being	Girls survey	STS	Self-esteem: 1.47 out of 3.00 mean score 60.94% of girls categorised as having a high score[87]	3320 83% of 4000 - cohorts 1 and 2	Yes
			STS	Self-confidence: 1.88 out of 3.00 mean score 75.35% of girls categorised as having a high score[88]	3320 83% of 4000 - cohorts 1 and 2	Yes
Main qualitative findings						
• Taboos regarding menstruation, specifically what girls can do while menstruating, are prevalent among adolescent girls. • Girls are believed to experience their sexual debut between age 12 and 15, while boys experience between 13 and 16. Condom use was the most frequently mentioned way to practice safe sex. • Girls most often get information on SRHR from female family members.						

[86] N=361
[87] N=361
[88] N=361

IO	IO indicator	Sampling and measuring technique used	Who collected the data?	Baseline level	Target for next evaluation point	Will IO indicator be used for next evaluation point? (Y/N)
• Girls expressed feeling a lack of bodily autonomy, specifically regarding unwanted touching and, in one instance, rape.						

Main findings

IO2.1 Number of girls with improved understanding of SRHR[89]

At baseline, girls were asked items on the girls' survey about their knowledge of sexual and reproductive health topics, such as sexually transmitted diseases, examples of SRHR and practices around SRHR topics. Items were reviewed and revised by TfaC to ensure alignment with the SRHR curriculum they will deliver over the life of the project. To report on IO2.1, an index of 18 items with a maximum score of 18.00 was created from a subset of SRHR items (see Supplemental Table 37 for list of items). The overall mean score on the SRHR understanding index at baseline was 4.06.

To further analyse SRHR understanding, girls were grouped in to 2 categories, where the median score, 4.00, served as the cut-off point. Girls who scored 4.00 or higher were categorised as 'high scores' (or above the median). Girls who scored less than the median of 4.00 were categorised as 'low scores' (or below the median). The proportions of girls in high and low score categories by subgroups and barriers are presented in Supplemental Table 18. The overall proportions in the population at baseline serves as a reference point against which proportions of girls in the high SRHR knowledge group can be compared by subgroup.

As expected, based on the cut-off score at the median, just about half (51.80%) of the 361 girls who reported on SRHR knowledge in the girls survey received high scores on the index. Age appeared to be an important factor in girls' SRHR knowledge. Only 10.53% of girls aged 10 through 11 had high scores, compared with 76.77% of girls aged 17 through 19. Girls in the married and caregiver subgroup and in the head of household subgroups also had high proportions of high scores (91.43% and 90.91%, respectively). Additionally, girls who faced bullying were more likely to have a low score on SRHR knowledge than were girls who did not.

Supplemental Table 18: IO2.1 SRHR understanding results by subgroup and barrier (median of 4.00 out of 18.00)

Category	N	Score	Proportion of total
All girls	**361**	**Low score**	**48.20%**
		High score	**51.80%**
Lilongwe	75	Low score	45.00%
		High score	55.00%
Dedza	180	Low score	49.33%
		High score	50.67%

[89] At baseline, TfaC also conducted evaluations of girls' SRHR knowledge, attitude and skills and completed a condom demonstration scoring card. Based on the items included in TfaC's tools, the average score of participants was 39.9%, which reflects low SRHR knowledge, attitudes and skills. On the condom demonstration scoring card, no participants were able to complete all steps correctly; the average percent score of successfully completed steps was 6% for male condom use and 1% for female condom use. See Annex 16: TfaC baseline results summary.

Category	N	Score	Proportion of total
Mchinji	106	Low score	52.83%
		High score	47.17%
Age 10–11	57	Low score	89.47%
		High score	10.53%
Age 12–16	205	Low score	48.78%
		High score	51.22%
Age 17–19	99	Low score	23.23%
		High score	76.77%
Subgroup			
Married and caregiver	70	Low score	8.57%
		High score	91.43%
Orphaned	95	Low score	43.16%
		High score	56.84%
Head of household	11	Low score	9.09%
		High score	90.91%
High chore burden	174	Low score	41.38%
		High score	58.62%
Girls with disabilities	125	Low score	48.80%
		High score	51.20%
Barrier			
Bullying	27	Low score	70.37%
		High score	29.63%
School cost	329	Low score	48.33%
		High score	51.67%
Parent support	67	Low score	50.75%
		High score	49.25%
Menstruation	108	Low score	53.70%
		High score	46.30%
Food insecurity or hunger	234	Low score	50.00%
		High score	50.00%
School safety	102	Low score	53.92%
		High score	46.08%

A regression model was conducted to understand the relative influence of multiple factors—girls' observable characteristics, marginalisation subgroups and barriers—on their level of SRHR understanding. Results are provided in Supplemental Table 19. Specific barriers were selected to include in the model based on relationships observed between variables—namely age, district, married and caregiver status, food insecurity and hunger.[90] The model showed that being in the 12 through 16 age group, the 17 through 19 age group and being a married caregiver were significant predictors of girls' SRHR understanding scores. When controlling for district, being married, a caregiver and experiencing food insecurity and hunger as a barrier (see Section 4.4 Quantitative evaluation methodology for explanation on inclusion and exclusion of variables). Specifically, if a girl is 12 through 16 years old, her SRHR score is likely to be 1.73 points higher than a girl who is 10 through 11 years old. Similarly, if a girl is 17 through 19 years old, her SRHR score is likely to be 2.30 points higher than if she is 10 through 11 years old. If a girl is married and a caregiver, her SRHR understanding score is likely to be 2.56 points higher than a girl who

[90] F (5, 355) =44.61, p<0.00.

is not, controlling for age, district and hunger status. Hunger, as a proxy measure of extreme poverty in the population, was not a statistically significant predictor of SRHR understanding.

Supplemental Table 19: IO2.1 SRHR understanding analytical model results

Category	Coefficient	Standard error	95% Confidence interval	
			Minimum	Maximum
Age 12–16	1.73***	0.28	1.19	2.28
Age 17–19	2.30***	0.33	1.66	2.95
Mchinji	-0.26	0.22	-0.70	0.18
Married and caregiver	2.56**	0.27	2.03	3.10
Food insecurity or hunger	-0.10	0.20	-0.50	0.30
Constant[91]	2.09	0.28	1.55	2.64

Note: One asterisk (*) denotes differences between groups that are statistically significant at $p < 0.05$, 2 asterisks (**) denote differences that are statistically significant at $p < 0.01$, and 3 asterisks (***) denote differences that are statistically significant at $p < 0.001$.

Qualitative data on IO2.1 were sourced from FGDs with adolescent girls and with adolescent boys. Findings from adolescent girls were grouped into several themes—menstruation, sexual activity, safe sex and contraception and other.

FGD findings on menstruation indicate a diverse terminology used by girls to talk about their period, including *kusamba*, *akusamba, tili kumwezi, apita ku mdima* and *kwabewera aleondo*. Girls mentioned that they learn about menstruation from female relatives, namely mothers, grandmothers and older sisters. They also said that they could not or do not talk about menstruation with male relatives—including brothers and fathers—other boys, small children and girls who have not started menstruating. When asked about activities that they should not or cannot do while menstruating, respondents in 4 of the 6 FGDs said that girls should not have sex during a monthly period. Several other taboos emerged from FGDs: girls should not use salt or collect clay while menstruating, nor should they take *panado* tablets, play ball with boys or greet people.[92]

Findings from FGDs on sexual activity indicated that girls believed that most or all their peers were sexually active—with the exception of younger adolescent girls. Girls discussed the general age of sexual debut as 12 through 15 years for girls and 13 through 16 for boys. The youngest age of sexual debut mentioned for girls was 5. Within several FGDs, girls said that the onset of menstruation is an indicator that a girl is ready for sex and a boyfriend.

When asked about community reactions to adolescent girls' pregnancies, several responses emerged. In FGDs with married girls, respondents mentioned that a pregnancy is an exciting event, although it sometimes elicits health concerns for younger girls. Girls expressed that community members think pregnancies come as a result of adolescent girls' want and do not blame the boys. In FGDs with caregivers, respondents said that members of the community refer to adolescent pregnant girls as prostitutes. Respondents in the FGD with orphans said that parents sometimes force their daughters out of the house or beat them if they become pregnant because the pregnancy brings shame to a family.

Girls were asked what safe sex means to young people their age. In nearly all FGDs, girls said that they consider condom use as safe sex.[93] Preventing or protecting oneself from pregnancy

[91] The constant, or intercept, is the average score for the reference group. In this case, the reference group is girls who are 10 or 11 years old from Dedza or Lilongwe who are not married caregivers and whose households have not gone to bed hungry 10 or more days in the last year.
[92] Panado (Paracetamol) is a non-opioid painkiller.
[93] The exception was for an FGD with girls aged 10 through 14, who said they did not know about safe sex.

and diseases was mentioned in several of the discussions. Only in the FGD with caregivers did respondents mention a method of contraception other than condoms—Depo-Provera, a contraceptive injection. Across FGDs, responses were inconsistent on who in a relationship has the responsibility for ensuring safe sex. Some girls said it was the boys' responsibility, some said it was the girl's and others said it was a shared responsibility. When asked how they learn about relationships, sex and contraception, most girls responded that they learn from their mothers, friends, grandmothers, *alangizi* (advisers), church counsellors, schools and medical staff, such as the mobile clinic Banja La Mtsogolo.

An emergent theme across FGDs with adolescent girls was their lack of bodily autonomy. Specifically, there were a number of mentions to touching of or violence against girls' bodies by young men and boys without girls' consent. When asked about the difference between boys and girls at their age, girls in one FGD said that boys are 'on top of the world' and enjoy touching girls' breasts and waists. In another FGD, respondents said that if a girl refuses a boy's advances, he will beat her. One respondent said that when a girl is passing a deserted place with a boy and she refuses his advances, she can be raped.

Responses from FGDs with adolescent boys indicated differing perceptions of gender roles in social and sexual relationships as well as persistent gender norms that impact these relationships. Across FGDs, boys described how girls and boys behave differently when with peers of the same gender. They shared that boys cannot discuss sexual relationships or family planning with girls or women, just with other men. Several boys described that girls ridicule or tease boys about sexual issues. While a respondent in the older boys' FGD said that the culture of the community favours girls and provides them with more support and counselling, several younger boys said that girls receive unwanted touch and are punished and beaten—especially when they become pregnant. Responses were mixed on which partner in a sexual relationship is responsible for contraception and protection. In one FGD, older boys agreed that it is the responsibility of the man because he 'knows what is good or bad', while in a different FGD, older boys said that it is a woman's responsibility not to get pregnant. When asked what girls should not or cannot do when they are menstruating, boys from across FGDs said that girls should not have sex, do too much work or be in class.

IO2.2 Number of girls with improved self-esteem, self-confidence and well-being[94]

Baseline values for IO2.2 were captured through the girls' survey using 2 different indices: self-esteem, which has 10 items, and self-confidence, which has 4 items (see Supplemental Table 38 for list of items).[95] At baseline, measures for well-being were not captured.[96] Both indices were standardised to a zero through three scale. The overall mean scores at baseline were 1.47 for self-esteem and 1.88 for self-confidence. Girls' scores were then categorised as high and low. High self-esteem scores were defined as scores above 1.50, the median of the self-esteem index. High self-confidence scores were defined as scores above 1.50, the median of the self-confidence index. As with the index for IO2.1, this imposed cut point provides a reference point against which proportions of girls in high and low groups can be compared by subgroup.

The proportions of girls in high and low score categories by subgroups and barriers are presented in Supplemental Table 20. A majority (60.94%) of girls had high self-esteem scores. A larger

[94] TfaC conducted a self-confidence observational assessment at baseline. The average overall self-confidence score was 27%, indicating low levels of self-confidence among participants. See Annex 16: TfaC baseline results summary.
[95] Link, TfaC and STS agreed to remove well-being from the measurement scale during the baseline survey design. This was not updated in the most recently approved logframe.
[96] In conversations with TEAM Girl Malawi prior to the baseline, the team indicated that well-being would be removed from Indicator IO2.2. As a result, data for this domain was not captured at the baseline evaluation point.

proportion of girls (75.35%) had high self-confidence scores. The proportions of girls in Mchinji in the high score category was smaller than for all girls on both the self-esteem and self-confidence indices—49.06% and 66.98%, respectively. The proportion of girls aged 17 through 19 in the high score categories were higher than for all girls—72.73% and 86.87%, respectively. The proportion of girls in the married and caregiver subgroup with high self-confidence scores was higher than for all girls—94.29% compared with 75.35%. While the proportion of girls who reported experiencing bullying as a barrier and had high scores was slightly larger than all girls on the self-esteem scale; however, it was notably lower on the self-confidence scale. Specifically, 51.85% of girls who experienced bullying had high self-confidence scores, compared with 75.35% of all girls.

Supplemental Table 20: IO2.2 Self-esteem and self-confidence results by subgroup and barrier (median of 1.50 out of 3.00)

Category	N	Score	Proportion of total (Self-esteem)	Proportion of total (Self-confidence)
All girls	**361**	**Low score**	**39.06%**	**24.65%**
		High score	**60.94%**	**75.35%**
Lilongwe	75	Low score	30.67%	26.67%
		High score	69.33%	73.33%
Dedza	180	Low score	35.56%	18.89%
		High score	64.44%	81.11%
Mchinji	106	Low score	50.94%	33.02%
		High score	49.06%	66.98%
Age 10–11	57	Low score	36.84%	33.33%
		High score	63.16%	66.67%
Age 12–16	205	Low score	45.37%	27.80%
		High score	54.63%	72.20%
Age 17–19	99	Low score	27.27%	13.13%
		High score	72.73%	86.87%
Subgroup				
Married and caregiver	70	Low score	38.57%	5.71%
		High score	61.43%	94.29%
Orphaned	95	Low score	43.16%	17.89%
		High score	56.84%	82.11%
Head of household	11	Low score	45.45%	18.18%
		High score	54.55%	81.82%
High chore burden	174	Low score	39.66%	16.67%
		High score	60.34%	83.33%
Girls with disabilities	125	Low score	40.00%	28.80%
		High score	60.00%	71.20%
Barrier				
Bullying	27	Low score	33.33%	48.15%
		High score	66.67%	51.85%
School cost	329	Low score	37.69%	25.84%
		High score	62.31%	74.16%
Parent support	67	Low score	26.87%	11.94%
		High score	73.13%	88.06%
Menstruation	108	Low score	36.11%	19.44%
		High score	63.89%	80.56%
Food insecurity or hunger	234	Low score	39.74%	24.79%
		High score	60.26%	75.21%

Category	N	Score	Proportion of total (Self-esteem)	Proportion of total (Self-confidence)
School safety	102	Low score	29.41%	26.47%
		High score	70.59%	73.53%

A regression model was conducted to determine the relative predictive influence of selected factors on girls' self-esteem and self-confidence. The selected factors include age, district, being married and a caregiver and experiencing food insecurity or hunger (see Section 4.4 Quantitative evaluation methodology for explanation on inclusion and exclusion of variables).[97] Results for self-esteem are presented in Supplemental Table 21 and in Supplemental Table 22 for self-confidence. The only statistically significant predictor on self-esteem was district. Specifically, girls from Mchinji are likely to score 0.15 points lower on the self-esteem scale than girls from Lilongwe or Dedza. Age group, being married and a caregiver and experiencing food insecurity or hunger were not significant predictors of girls' self-esteem scores. While not included as one of the core predictors of the regression models, bullying was a predictor of girls' self-confidence. Girls who reported higher levels of bullying had lower self-confidence.

Supplemental Table 21: IO2.2 Self-esteem analytical model results

Category	Coefficient	Standard error	95% Confidence interval	
			Min.	Max.
Age 12–16	-0.04	0.28	1.19	2.28
Age 17–19	0.09	0.33	1.66	2.95
Mchinji	-0.15*	0.22	-0.70	0.18
Married and caregiver	0.00	0.27	2.03	3.10
Food insecurity or hunger	0.00	0.20	-0.50	0.30
Constant[98]	1.51	0.07	1.38	1.65

Note: One asterisk (*) denotes differences between groups that are statistically significant at $p < 0.05$. Two asterisks (**) denote differences that are statistically significant at $p < 0.01$ and 3 asterisks (***) denote differences that are statistically significant at $p < 0.001$.

Findings from the initial analysis did not show age as a statistically significant predictor of self-confidence scores. Instead, starting menstruation was included in the model along with district, being married and a caregiver and experiencing food insecurity or hunger (see Section 4.4 for explanation on inclusion and exclusion of variables).[99, 100] Results showed that having started menstruation was a statistically significant predictor, as was being from Mchinji and being married and a caregiver, when controlling for food insecurity and hunger as a barrier. Specifically, girls who have started menstruation are likely to score 0.31 points higher on the self-confidence scale than girls who have not, controlling for other factors in the model. Girls who are married and caregivers are likely to score 0.37 points higher on the self-confidence scale, controlling for other factors in the model. Conversely, girls from Mchinji are likely to score 0.28 points lower on the scale, controlling for other factors in the model. Experiencing food insecurity or hunger as a barrier was not significant predictors of girls' self-confidence scores.

[97] F (5, 355) =3.61, p<0.01
[98] The constant, or intercept, is the average score for the reference group. In this case, the reference group is girls who are 10 or 11 years old from Dedza or Lilongwe who are not married caregivers and whose households have not gone to bed hungry 10 or more days in the last year.
[99] Age and menstruation were not included in the same model as there was high correlation between these variables. Age was not a statistically significant predictor in this model, but menstruation was.
[100] F (4, 351) =9.49, p<0.0001

Supplemental Table 22: IO2.2 Self-confidence analytical model results

Category	Coefficient	Standard error	95% Confidence interval	
			Min.	Max.
Started menstruation	0.31**	0.10	0.12	0.51
Mchinji	-0.28**	0.10	-0.47	-0.09
Married and caregiver	0.37**	0.11	0.15	0.60
Food insecurity or hunger	-0.02	0.09	-0.20	0.17
Constant[101]	1.69	0.10	1.48	1.89

Note: One asterisk (*) denotes differences between groups that are statistically significant at p < 0.05. Two asterisks (**) denote differences that are statistically significant at p < 0.01 and 3 asterisks (***) denote differences that are statistically significant at p < 0.001.

Girls were not asked direct questions regarding their self-esteem and self-confidence in FGDs, although themes mentioned in Supplemental Table 17 and Supplemental Table 23 may contribute negatively to their self-perception under these constructs.

Reflections

Girls' baseline SRHR understanding was low, indicating there is sufficient room for growth before the next evaluation point. Girls' baseline scores on the self-esteem index also show room for growth (mean score of 1.43 out of 3.00). Conversely, girls' baseline self-confidence mean scores were higher (1.88 out of 3.00), indicating that there may be less room for growth. There is evidence of a relationship between girls' SRHR understanding and their learning outcomes as well as between girls' self-confidence and their learning outcomes (see Supplemental Table 7), which supports their inclusion in the logframe and project's ToC.

Indicator IO2.2 currently includes the term 'well-being' but no items were available to report against this aspect of the indicator. The indicator should be revised to remove this term as it is not expected to be added to tools in subsequent timepoints.

Specifically, STS recommends the following updates to indicators under IO2:

- IO2.2: Number of girls with improved self-esteem and self-confidence

At the next evaluation point, new tools or items will be introduced to track indicators under IO2. These include:

For IO2.2:

- Self-confidence observation checklist (developed and piloted by TfaC)
- Additional girls survey items

Qualitative data for all IO2 indicators will be captured from adolescent girls and boys, mothers' group members, caregivers, community leaders and CBE facilitators. Additional focus will be placed on exploring self-esteem and self-confidence in FGDs with adolescent girls and boys.

Targets

STS proposes the following targets for year 3:

- IO2.1 Number of girls with improved understanding of SRHR: 75% of enrolled girls

[101] The constant, or intercept, is the average score for the reference group. In this case, the reference group is girls who have not started menstruation from Dedza or Lilongwe who are not married caregivers and whose households have not gone to bed hungry 10 or more days in the last year.

- IO2.2 Number of girls with improved self-esteem and self-confidence: 70% of girls have improved self-esteem, 65% of girls have improved self-confidence

Project Checks on Intermediate Outcomes

Ensure that the IO analysis reflects the links between different levels in the logframe and informs the validity of the Theory of Change. This includes checking whether the EE (?) have:

- Measured and analysed all IO indicators presented in logframe.
- Disaggregated the data according to the logframe.
- Used both the qualitative and quantitative analysis stated in the logframe.
- Related the IO analysis to the analysis of Outcomes.

Project confirmation that all the IO analysis has been covered.

IO3: Improvement in quality of education at CBE centres and primary schools

TEAM Girl Malawi's third IO is improvement in the quality of education at CBEs and primary schools. Specifically, the project ToC assumes that improved quality of education is a prerequisite for better learning, transition and sustainability outcomes for marginalised girls.

Formal learning instruction had not yet begun at CBE centres at baseline and given that the project had not yet begun intervening in primary schools, the focus of data collection for this indicator was to identify perceived challenges to learning in the school environment. Because girls and boys had not yet begun formal sessions at the CBE and were out of school, their responses were based on a perception of anticipated challenges given their previous experiences in the school system and likely not barriers directly experienced at CBEs. TEAM Girl Malawi should use this information to help inform the structure and focus of interventions for CBE facilitators and primary schools.

Table 25: IO3 Improvement in quality of education at CBE centres and primary schools' indicators

IO	IO indicator	Sampling and measuring technique used	Who collected the data?	Baseline level	Target for next evaluation point	Will IO indicator be used for next evaluation point? (Y/N)
IO3: Improvement in quality of education at CBE Centres and Primary Schools	IO3.1: Number of CBE facilitators practising gender responsive pedagogy and inclusive and child-centred literacy and numeracy teaching methodologies	Classroom observation CBE facilitator survey Girls survey	NA at baseline	0	160	Yes
IO3: Improvement in quality of education at CBE Centres	IO3.2: Number of head teachers, teachers and CBE facilitators reporting positive	CBE facilitator survey	NA at baseline	0	TBC	Yes

IO	IO indicator	Sampling and measuring technique used	Who collected the data?	Baseline level	Target for next evaluation point	Will IO indicator be used for next evaluation point? (Y/N)
and Primary Schools	changes in gender perceptions and gender-sensitive teaching	Girls survey				
IO3: Improvement in quality of education at CBEs and primary schools	IO3.3: Number of schools registering improvements against the National Education Standards	Head teacher survey	NA at baseline	0	TBC	TBC
Main qualitative findings						
• Potential challenges to the quality of learning environments at CBEs, include relationships between facilitators, teachers and learners—concerns about animosity or poor relationships were frequently mentioned during qualitative research, as were fears of corporal punishment. • School safety was a major concern of respondents, who mentioned fears of bullying, fighting and physical violence at schools as a barrier to a quality learning environment.						

Main findings

Because learning sessions had not yet begun, quantitative baseline data for indicators IO3.1, IO3.2 and IO3.3 were not collected. Anticipated learning environment barriers and challenges that were mentioned by adolescent girls and boys during FGDs are described in Supplemental Table 23. These are perceived barriers and challenges based on the beneficiaries' previous experiences in the public schooling system.

Supplemental Table 23: Learning environment barriers by subgroup from qualitative data

Barrier	Subgroup characteristic
Facilitator or teacher absenteeism	Adolescent girls aged 15–19 Caregivers Adolescent boys aged 15–19
Conflict with facilitator or teacher, including disliking or hating them or vice versa	Adolescent girls aged 15–19 Married Orphaned
Physical violence perpetrated by facilitator or teacher against learner	Adolescent girls aged 10–14
Physical or emotional abuse among learners, including being bullied or tortured by peers, beaten by peers, fighting with peers or hating peers	Adolescent girls aged 10–14 Orphaned Married Caregivers Girls with disabilities
Having school materials stolen in classroom	Adolescent girls aged 10–14 Girls with disabilities

As evidenced through FGDs and quantitative surveys, a potential challenge in the learning environments may be relationships between the facilitators and teachers with the learners.

Additionally, the questions asked in surveys regarding bullying included whether girls experienced bullying from peers or teachers. Bullying emerged as a significant barrier to girls learning and IOs. Concerns regarding animosity or poor relationships between educators and learners were frequently mentioned, as were fears of corporal punishment. Respondents were concerned about the safety of the learning environment. Bullying, fighting and physical violence were all mentioned as learning environment challenges.

Reflections

Because learning sessions had not yet started at the time of data collection, data collection for IO3 focused primarily on identification of learning environment barriers and challenges through qualitative methods. The findings under IO3 at baseline should be used to provide formative feedback to the project to reduce barriers that girls and boys may confront in the coming years of the project.

The project should consider simplifying indicator IO3.2—number of head teachers, teachers and CBE facilitators reporting positive changes in gender perceptions and gender-sensitive teaching—to capture data from only CBE facilitators. As the project's intervention may be less involved in primary schools, it is likely that the most statistically significant changes in gender perceptions and gender-sensitive teaching will be experienced by CBE facilitators. Further, changing the indicator respondent to solely CBE facilitators will reduce the cost of quantitative data collection under this indicator.

The project should also critically evaluate whether planned interventions at the primary school level may lead to intended improvements under IO3.3—number of schools registering improvements against the National Education Standards.

Specifically, STS recommends the following updates to indicators under IO3:

- IO3.2: Percentage of head teachers, teachers and CBE facilitators that report positive changes in gender perceptions and gender-sensitive teaching

At the next evaluation point, several new tools or items will be introduced to track indicators under IO3. These include:

For IO3.1:

- CBE facilitator surveys
- CBE classroom observations
- Additional girls survey items

For IO3.2:

- CBE facilitator surveys
- Additional girls survey items

For IO3.3:

- Head teacher surveys

Qualitative data for IO3 indicators will be captured from adolescent girls and boys, CBE facilitators, head teachers and teachers.

Targets

STS proposes the following targets for year 3:

- IO3.1 Number of CBE facilitators practising gender responsive pedagogy & inclusive and child-centred literacy and numeracy teaching methodologies: 72 (90% of CBE facilitators – out of a total of 80 facilitators in cohort 1)
- IO3.2 Percentage of head teachers, teachers and CBE facilitators that report positive changes in gender perceptions and gender-sensitive teaching: 40 percent
- IO3.3 Number of schools registering improvements against the National Education Standards: 50% of schools

Project confirmation that all the IO analysis has been covered.

IO4: Improvement in community members' understanding and use of support mechanisms for marginalised girls

TEAM Girl Malawi's fourth IO is improvement in community members' understanding and use of support mechanisms for marginalised girls. Specifically, the project ToC assumes that improved understanding and use of these mechanisms are prerequisites for better learning, transition and sustainability outcomes for marginalised girls.

IO4 indicators and relevant baseline information are detailed in Table 26. Indicator IO4.1 is zero at baseline, as girls were out of school and had not yet started learning sessions in CBEs. Indicators IO4.2 and IO4.3 are reported as mean scores at the household level. At midline, the proportion of households with improved scores over baseline will be reported for both indicators.

Table 26: IO4 Improvement in community members' understanding and use of support mechanisms for marginalised girls' indicators

IO	IO indicator	Sampling and measuring technique used	Who collected the data?	Baseline level	Target for next evaluation point	Will IO indicator be used for next evaluation point? (Y/N)
IO4: Improvement in community members' understanding and use of support mechanisms for marginalised girls	IO4.1: Number of girls who report feeling safe at CBEs and primary schools	Girls survey	NA at baseline	0	TBC	Yes
IO4: Improvement in community members' understanding and use of support	IO4.2: Improved community support for SRHR and CP	Girls survey Household survey	STS	SRHR – 2.18 out of 4.00 mean score[102]	SRHR – TBC	Yes

[102] SRHR support mean scores were calculated for those who answered 25% or more of the items in the index (N=371).

IO	IO indicator	Sampling and measuring technique used	Who collected the data?	Baseline level	Target for next evaluation point	Will IO indicator be used for next evaluation point? (Y/N)
mechanisms for marginalised girls				CP – 2.60 out of 3.00 mean score[103]	CP – TBC	
IO4: Improvement in community members' understanding and use of support mechanisms for marginalised girls	IO4.3: Improved community support for girls' education through CBEs and primary school	Girls survey Household survey	STS	10.05 out of 15.00 mean score[104]	TBC	Yes
Main qualitative findings						
• Key in-school safety concerns for girls include bullying, fighting and physical violence. • Community leaders indicate some degree of existing awareness of and support for SRHR, primarily through nongovernmental organisations, community-based organisations, government health workers and mothers' groups; openness to discussing issues varies across different communities. • Community leaders and mothers' group members expressed strong personal and community support for marginalised girls' education, although barriers to learning and transition exist in communities.						

Main findings

IO4.1 Number of girls who report feeling safe at CBEs and primary schools

As formal learning sessions had not yet begun at CBE and beneficiaries were out of school at baseline, no baseline results are reported for IO4.1. There were, however, quantitative and qualitative indications of safety concerns travelling to and from learning centres, though the response rate on the girls' survey to safety items was low. Specifically, when respondents were asked if they felt safe travelling to and from school, 46.15% reported that they did not (n=18). Additionally, 23.08% of girls reported feeling unsafe at school (n=9) (Annex 15: Additional tables Supplemental Table 35). Of the 353 caregivers responding to the household survey, 24.08% reported that it was fairly or very unsafe for girls to travel to schools in their area. Of the 316 caregivers who reported that their girl was not enrolled in primary school, 11.71% said it was because it is unsafe to travel to or from school in that area, while 8.54% reported that it was because their girl is not safe at school (Supplemental Table 40).

[103] CP support mean scores were calculated for those who answered 25% or more of the items in the index (N=371).
[104] Girls' education support mean scores were calculated for those who answered 25% or more of the items in the index (N=370).

Qualitative data corroborated the safety concerns expressed by girls and caregivers in surveys (see Supplemental Table 17 for individual-, household- and community-level safety concerns, and Supplemental Table 23 for school-level safety concerns).

IO4.2 Improved community support for SRHR and child protection

Results for IO4.2 at baseline are reported as mean scores on a SRHR support index and CP support index. The community support for SRHR index was created from items in the household survey. The CP index was created from items on the girls and household surveys that were combined into a single score (see Supplemental Table 39 for index items). Items were reviewed and revised by TfaC to ensure alignment with the SRHR and CP curriculum they will deliver over the life of the project.

Mean scores for the SRHR support index are presented in Supplemental Table 24. While the reliability of this index was notably low, the index was constructed using the items that adhered to the theory underlying the design of the instruments. Further analysis at the item level revealed that the low reliability may, in fact, be attributed to discrepancies in responses that reflect the discrepancies in perceptions around SRHR.[105] Other items that contributed to the lower reliability had very poor differentiation between respondents, making items less useful in the overall composite score.[106] However, as these items were included based on the project's underlying approach to increasing community support for SRHR, the index utilises all these items. The index construction is based on theory, and the low reliability is primarily due to the underlying patterns of responses on items. As a result, the use of the index is warranted at baseline

Across households, the mean score on the SRHR support index was 2.18, out of a maximum 4.00. Mean scores for households with girls who were married and caregivers were statistically significantly higher than all other households, 2.32 compared with 2.15, respectively. Households that faced a food insecurity or hunger barrier had statistically significantly lower mean scores than households that did not face this barrier, 2.14 compared with 2.27, respectively.

Supplemental Table 24: IO4.2 SRHR support mean scores by subgroup and barrier

Category	Disaggregation	N/n	Mean score (out of 4.00)
All households	**NA**	**348**	**2.18**
District	Dedza	174	2.17
	Lilongwe	72	2.24
	Mchinji	102	2.15
Age	Age 10–11	56	2.10
	Age 12–16	204	2.20
	Age 17–19	88	2.20
Subgroup			
Married and caregiver**	Not in subgroup	282	2.15
	In subgroup	66	2.32
Orphaned	Not in subgroup	251	2.19

[105] For example, respondents who had high overall scores on the index tended to disagree a lot with the statement 'a wife can be beaten up if she does not listen to or obey her husband' (54.00%) but so did respondents who had low overall scores (40.80%). Similarly, the proportions who agreed 'a lot' with the statement were not more likely to respond in a particular way on the remaining items in the index—46.00% of respondents with low overall score on the index agreed a lot with the statement and 45.40% of respondents with high overall score on the index also agreed a lot with the statement. An item that was a clear differentiator between those who were low and high on the index was, 'I believe that girls have the right to go to school while pregnant'. The majority of those with high scores on the index said this statement was completely true (82.60%) while the majority of those with low scores on the index said the statement was not true (85.60%).

[106] One example includes, 'I believe that girls have the right to go back to school after they have children', which was overwhelmingly found to be 'completely true' by the majority of respondents.

Category	Disaggregation	N/n	Mean score (out of 4.00)
	In subgroup	97	2.16
Head of household	Not in subgroup	342	2.18
	In subgroup	6	2.25
High chore burden	Not in subgroup	175	2.16
	In subgroup	173	2.20
Girls with disabilities	Not in subgroup	229	2.21
	In subgroup	119	2.12
Barrier			
Bullying	Does not face barrier	319	2.18
	Faces barrier	29	2.16
School cost	Does not face barrier	32	2.19
	Faces barrier	316	2.18
Parent support	Does not face barrier	278	2.19
	Faces barrier	70	2.12
Menstruation	Does not face barrier	239	2.17
	Faces barrier	109	2.20
Food insecurity or hunger**	Does not face barrier	112	2.27
	Faces barrier	236	2.14
School safety	Does not face barrier	248	2.20
	Faces barrier	100	2.14

Note: Two asterisks (**) denote differences that are statistically significant at p < 0.01.

STS conducted a regression model to determine the relative predictive influence of households' SRHR support. Findings are presented in Supplemental Table 25. Results indicate that experiencing food insecurity or hunger as a barrier and having a girl who is married and a caregiver are statistically significant predictors of scores on the SRHR support index, controlling for girls' age and district (see Section 4.4 Quantitative evaluation methodology for explanation on inclusion and exclusion of variables).[107] Households with girls who are married and caregivers are likely to have a 0.20 point higher score on the index than households without girls who are married and caregivers. Households that experience food insecurity or hunger as a barrier are likely to score 0.14 points lower on the SRHR support scale. Age and district were not statistically significant predictors of household scores on the Support for SRHR index.

Supplemental Table 25: IO4.2 SRHR support analytical model results

Category	Coefficient	Standard error	95% Confidence interval	
			Min.	Max.
Age 12–16	0.09	0.06	-0.03	0.21
Age 17–19	0.01	0.08	-0.14	0.16
Mchinji	-0.03	0.05	-0.13	0.07
Married and caregiver	0.20**	0.06	0.08	0.32
Food insecurity or hunger	-0.14**	0.05	-0.24	-0.05
Constant[108]	2.19	0.06	2.07	2.32

Note: Two asterisks (**) denote differences that are statistically significant at p < 0.01.

[107] F (5, 342) =4.34, p<0.001

[108] The constant, or intercept, is the average score for the reference group. In this case, the reference group is girls who are 10 or 11 years old from Dedza or Lilongwe who are not married caregivers and whose households have not gone to bed hungry 10 or more days in the last year.

KIIs with community leaders indicated some degree of existing community support for SRHR as well as mechanisms for accessing information on SRHR, although perceptions appeared to vary across communities. One community leader described that nongovernmental organisations and community-based organisations help girls to understand SRHR, as do government health workers. Mothers' groups were also mentioned as key sources of information and awareness-raising on SRHR. The same community leader said that there is a shift in the community towards more open discussion of SRHR. One community leader estimated that 90% of people in the community were supportive of strengthening girls' SRHR, and that some organisations and mothers' groups incorporated drama and dances to better engage and teach girls about SRHR. A different community leader said that the main source of information on SRHR in the community was school, health workers and mothers' groups, and in the latter, that abstinence was a key message. This community leader noted that parents in the community do not discuss SRHR issues with adolescent girls.

Mean scores for the CP support index are presented in Supplemental Table 26. The overall household mean score on the CP support index was 2.60 on a 3.00-point scale. Households with girls aged 10–11 had statistically significantly lower mean scores (2.24) than did all other households. Households with girls aged 17–19 had statistically significantly higher mean scores (2.73) than did all other households. Further, mean scores for households with girls who were married and caregivers were statistically significantly greater than all other households, 2.81 compared with 2.55, respectively.

Supplemental Table 26: IO4.2 Child protection support mean scores by subgroup and barrier

Category	Disaggregation	N/n	Mean score (out of 3.00)
All households	**NA**	**371**	**2.60**
District	Dedza	188	2.64
	Lilongwe	77	2.65
	Mchinji	106	2.52
Age	Age 10–11***	58	2.24
	Age 12–16	211	2.64
	Age 17–19*	102	2.73
Married and caregiver**	Not in subgroup	298	2.55
	In subgroup	73	2.81
Orphaned	Not in subgroup	273	2.59
	In subgroup	98	2.64
Head of household	Not in subgroup	360	2.61
	In subgroup	11	2.66
High chore burden	Not in subgroup	192	2.55
	In subgroup	179	2.66
Girls with disabilities	Not in subgroup	241	2.61
	In subgroup	130	2.59
Bullying	Does not face barrier	342	2.62
	Faces barrier	29	2.43
School cost	Does not face barrier	37	2.79
	Faces barrier	334	2.58
Parent support	Does not face barrier	301	2.63

Category	Disaggregation	N/n	Mean score (out of 3.00)
	Faces barrier	70	2.51
Menstruation	Does not face barrier	253	2.59
	Faces barrier	118	2.63
Food insecurity or hunger	Does not face barrier	135	2.58
	Faces barrier	236	2.62
School safety	Does not face barrier	268	2.60
	Faces barrier	103	2.61

Note: One asterisk (*) denotes differences between groups that are statistically significant at p < 0.05. Two asterisks (**) denote differences that are statistically significant at p < 0.01 and 3 asterisks (***) denote differences that are statistically significant at p < 0.001.

Results from the regression model for CP support are presented in Supplemental Table 27. Findings indicate that having a girl in the 12 through 16 or the 17 through 19 age groups was a statistically significant predictor of household scores on the index, controlling for district, having a girl who is married and a caregiver and experiencing food insecurity or hunger as a barrier (see Section 4.4 Quantitative evaluation methodology for explanation on inclusion and exclusion of variables).[109] Households with girls between 12 and 16 years old are likely to have CP support scores that are 0.41 points higher than households with girls aged 10 or 11. Similarly, households with girls aged 17 through 19 years old are likely to have CP scores 0.39 points higher than households with girls aged 10 or 11. District, having a girl who is married and a caregiver and experiencing food insecurity or hunger as a barrier were not statistically significant predictors of household CP support scores.

Supplemental Table 27: IO4.2 Child protection support analytical model results

Category	Coefficient	Standard error	95% Confidence interval	
			Min.	Max.
Age 12–16	0.41***	0.10	0.22	0.61
Age 17–19	0.39**	0.12	0.16	0.62
Mchinji	-0.14	0.08	-0.29	0.02
Married and caregiver	0.19	0.10	0.00	0.37
Food insecurity or hunger	0.05	0.07	-0.10	0.19
Constant[110]	2.23	0.10	2.04	2.43

Note: Two asterisks (**) denote differences that are statistically significant at p < 0.01 and 3 asterisks (***) denote differences that are statistically significant at p < 0.001.

CP support and mechanisms were explored in FGDs with mothers' groups and KIIs with community leaders. Specifically, CP was discussed in relation to prevalence and responses to physical and sexual violence against children and young people in their communities. Across communities, there appeared to be some mechanisms for supporting CP and some existing reporting and referral systems. Examples of resources for sexual abuse and rape mentioned included community police, village chiefs or heads and health services. Community leaders and one FGD participant mentioned a CP committee as a resource for the person experiencing violence. In cases of child abuse perpetrated by caregivers, one mothers' group member indicated that, when she witnesses physical violence such as slapping or hitting, she informs children to

[109] F (5, 365) =5.93, p<0.0001

[110] The constant, or intercept, is the average score for the reference group. In this case, the reference group is girls who are 10 or 11 years old from Dedza or Lilongwe who are not married caregivers and whose households have not gone to bed hungry 10 or more days in the last year.

report their caregivers. Mothers' group members highlighted that children who are orphaned or living with foster or stepparents are more vulnerable to abuse.

Mothers' group members in one FGD said that violence at school is reported to the School Management Committee and the mothers' group, who refer the case to the Ikata. One mothers' group member noted that the Ikata is effective and acts quickly on reports of school-based violence.

Mechanisms to address early marriage were described by respondents. These include village chiefs or heads making proclamations that anyone 14 years and younger who are married will be punished, must pay a goat and that there will be legal ramifications for parents involving the Ikata. Mothers' group members and community leaders noted that, although the prevalence of early marriage has decreased and the average age of marriage has increased, the issue is still common. Respondents noted that the perpetuation was likely due to contributing factors, such as poverty, peer pressure, early sexual debut, transactional sex and early pregnancy.

IO4.3 Improved community support for girls' education through CBEs and primary school

Results for IO4.3 are reported at baseline as a mean score on girls' education support index, created from items on the girls and household surveys. Index items from across the 2 surveys were combined into a household score. The 12 items that comprised the index were related to attitudes towards girls' education, gender norms and aspirations for girls after completing CBE (see Supplemental Table 40 for list of items).[111] The maximum score on the 12-item girls' education support index was 15.00.

Mean scores by subgroup and barrier are presented in Supplemental Table 28. The household mean score on the girls' education support index was 10.05 on a 15.00-point scale. Households in Dedza had statistically significantly lower mean scores (9.69) than all other households, while households in Lilongwe had significantly higher mean scores (11.23) than all other households. Mean scores for households with girls aged 10 or 11 were significantly higher than all other households (10.67). Households with girls aged 17 through 19 had significantly lower mean scores than all other households (9.33).

Further, households with girls who were married and caregivers had statistically significantly lower mean scores on the girls' education support index than did all other households, 9.67 compared with 10.15, respectively. Similarly, households with girls with disabilities had significantly lower mean scores than all other households, 9.61 compared with 10.30, respectively. Households with girls facing barriers related to menstruation had significantly lower mean scores than other households—9.56 compared with 10.28—while households facing food insecurity or hunger barriers had significantly higher mean scores than all other households—10.39 compared with 9.46.

Supplemental Table 28: IO4.3 Girls' education support mean scores by subgroup and barrier

Category	Disaggregation	N/n	Mean score (out of 15.00)
All households	**NA**	**370**	**10.05**
District	Dedza**	187	9.69
	Lilongwe***	77	11.23
	Mchinji	106	9.84

[111] One item on the household survey was split into four separate variables, as the item allowed multiple responses. See Supplemental Table 40.

Category	Disaggregation	N/n	Mean score (out of 15.00)
Age	Age 10–11*	58	10.67
	Age 12–16	210	10.23
	Age 17–19**	102	9.33
Married and caregiver**	Not in subgroup	298	10.15
	In subgroup	72	9.67
Orphaned	Not in subgroup	272	9.88
	In subgroup	98	10.55
Head of household	Not in subgroup	359	10.11
	In subgroup	11	8.27
High chore burden	Not in subgroup	192	10.17
	In subgroup	178	9.93
Girls with disabilities*	Not in subgroup	240	10.30
	In subgroup	130	9.61
Bullying	Does not face barrier	341	10.12
	Faces barrier	29	9.24
School cost	Does not face barrier	36	8.89
	Faces barrier	334	10.18
Parent support	Does not face barrier	301	10.11
	Faces barrier	69	9.80
Menstruation*	Does not face barrier	253	10.28
	Faces barrier	117	9.56
Food insecurity or hunger**	Does not face barrier	134	9.46
	Faces barrier	236	10.39
School safety	Does not face barrier	267	10.00
	Faces barrier	103	10.18

Note: One asterisk (*) denotes differences between groups that are statistically significant at $p < 0.05$. Two asterisks (**) denote differences that are statistically significant at $p < 0.01$ and 3 asterisks (***) denote differences that are statistically significant at $p < 0.001$.

Results from the regression model for girls' education support are presented in Supplemental Table 29. Age was not found to be a statistically significant predictor in the model and thus was excluded. Having a girl who had started menstruation was, however, significant and, as a result, was included in the analytical model (see Section 4.4 Quantitative evaluation methodology for explanation on inclusion and exclusion of variables).[112] Overall, starting menstruation, being from Mchinji and experiencing food insecurity or hunger were statistically significant predictors of girls' education support, controlling for caregiver and marital status.[113] Households with girls who have started menstruation are likely to score 0.83 points lower on the scale, and households from Mchinji are likely to score 0.59 points lower on the scale. Households that experience food insecurity or hunger as a barrier are likely to score 0.84 points higher on the scale than other households. Being in a household with a girl who is married and a caregiver was not a statistically significant predictor of girls' education support scores.

[112] Age and menstruation were not included in the same model as there was high correlation between these variables.
[113] $F(5, 351) = 36.08$, $p<0.001$

Supplemental Table 29: IO4.3 Girls' education support analytical model results

Category	Coefficient	Standard error	95% Confidence interval	
			Min.	Max.
Started menstruation	-0.83**	0.28	-1.39	-0.27
Mchinji	-0.59*	0.28	-1.14	-0.04
Married and caregiver	-0.35	0.33	-1.00	0.30
Food insecurity or hunger	0.84**	0.27	0.31	1.37
Constant[114]	10.41	0.30	9.83	10.99

Note: One asterisk (*) denotes differences between groups that are statistically significant at $p < 0.05$. Two asterisks (**) denote differences that are statistically significant at $p < 0.01$.

In KIIs and FGDs, community leaders and mothers' group members expressed strong personal and community support for marginalised girls' education. One community leader described that parents in the community had built improved water, sanitation and hygiene facilities at a school to ensure girls had a place to go for menstrual hygiene. Another community leader mentioned engagement and outreach with parents through community meetings, which resulted in improved attendance. Despite this support, adolescent girls and boys described existing household and community barriers to education. These include caregivers prioritising household or domestic work over girls' education, discouragement from attending school from relatives or husbands, lack of parental financial support to attend school, peer pressure and lack of safety on the way to and from school (see Supplemental Table 17).

Reflections

Because girls had not started learning sessions at CBE and were out of school at baseline, data to respond to indicator IO4.1 was not available. Instead, data captured looked whether girls and their caregivers believed lack of safety to be an impediment to attending CBE or primary school.

Baseline results for indicator IO4.2—2.18 out of 4.00 on SRHR support index and 2.60 out of 3.00 on CP support index—suggest potential challenges in showing growth over time. The mean score for the CP support index was notably high on the index, and both indices had relatively low reliability.[115] Further, there is evidence of a relationship between households' CP support score and girls' learning outcomes. Before the next evaluation point, mitigating steps will be taken to strengthen the reliability of IO4.2 indices and to allow for further growth over time. Results for indicator IO4.3 from baseline suggest that households have room to improve girls' education support over time.

The following updates to indicators, which will better articulate the intended respondent and change, are proposed:

- IO4.1: Percentage of girls who report feeling safe at CBEs and primary schools
- IO4.2: Percentage of households with improved support for SRHR and CP
- IO4.3: Percentage of households with improved support for girls' education through CBEs and primary schools

TEAM Girl Malawi should also confer with TfaC to confirm if SRHR support at the household level is an intended outcome of their implementation.

[114] The constant, or intercept, is the average score for the reference group. In this case, the reference group is girls who have not started menstruation from Dedza or Lilongwe who are not married caregivers and whose households have not gone to bed hungry 10 or more days in the last year.
[115] Additional items may be added and substituted for existing items, based on results obtained at the next evaluation point.

At the next evaluation point, new tools or items will be introduced to track indicators under IO4. These include:

For IO4.1:

- Additional girls survey items

For IO4.2:

- Additional household survey items

For IO4.3:

- Additional household survey items

Qualitative data for IO4 indicators will be captured from adolescent girls and boys, mother's group members, caregivers and community leaders.

Targets

STS proposes the following targets for year 3:

- IO4.1 Percentage of girls who report feeling safe at CBEs and primary schools: 65%
- IO4.2 Percentage of households with improved support for SRHR and CP: 70% for SRHR support, 60% for CP support
- IO4.3 Percentage of households with improved support for girls' education through CBEs and primary schools: 60%

Project confirmation that all the IO analysis has been covered.

IO5: Strengthened district and national leadership and engagement in marginalised adolescent girls' education

TEAM Girl Malawi's fifth IO is strengthened district and national leadership and engagement in marginalised adolescent girls' education. Specifically, the project ToC assumes that stronger governmental engagement in marginalised adolescent girls' education is a prerequisite for better learning, transition and sustainability outcomes for marginalised girls.

IO5 indicators and relevant baseline information are detailed in Table 27. Baseline data for IO5 was comprised of qualitative findings from KIIs and desk research completed during the GESI Analysis.

Table 27: IO5 Strengthened district and national leadership and engagement in marginalised adolescent girls' education indicators

IO	IO indicator	Sampling and measuring technique used	Who collected the data?	Baseline level	Target for next evaluation point	Will IO indicator be used for next evaluation point? (Y/N)
IO5: Strengthened district and national leadership	IO5.1: Number of national and district plans and policies that include measures	Count of policies using GESI Analysis	STS	10	TBC	Yes

IO	IO indicator	Sampling and measuring technique used	Who collected the data?	Baseline level	Target for next evaluation point	Will IO indicator be used for next evaluation point? (Y/N)
and engagement in marginalised adolescent girls' education	to strengthen primary schools to meet the needs of marginalised adolescent girls	KII data from district-level MOEST and MOGCDSW officials				
IO5: Strengthened district and national leadership and engagement in marginalised adolescent girls' education	IO5.2: Number of district education conferences that reference marginalised girls' education	District education conference minutes	NA at baseline	0	TBC	TBC
IO5: Strengthened district and national leadership and engagement in marginalised adolescent girls' education	IO5.3: National and district government staff demonstrate supportive attitudes to marginalised girls' education	Government official survey	NA at baseline	0	TBC	TBC
Main qualitative findings						
• Knowledge of plans and policies to strengthen marginalised girls' education varied across respondents, with some district-level officials saying they were not aware of national-level policies. • Government officials expressed support for marginalised girls' education but raised concerns over resources.						

Main findings

IO5.1: Number of national and district plans and policies that include measures to strengthen primary schools to meet the needs of marginalised adolescent girls

In KIIs, government officials were asked about their knowledge of plans and policies that include measures to strengthen primary schools to meet the needs of marginalised adolescent girls in order to respond to IO5.1. A list of the plans and policies relevant to marginalised adolescent girls are detailed in Supplemental Table 30. Some were mentioned in KIIs while others were referenced in the GESI Analysis. Description and awareness of policies was not consistent across

government respondents. During KIIs, one district-level MOEST respondent said he was not conversant with the National Girls' Education Strategy nor National Girls' Education Communication Strategy. Many respondents referenced policies by informal or partial names. In addition to the plans and policies directly related to primary schools, MOGCDSW officials referenced a number of other policies related to marginalised populations, CPs and gender equality.

Supplemental Table 30: Existing national and district plans and policies that reference marginalised adolescent girls

Policy	Source
Malawi National Girls' Education Strategy 2013+	GESI Analysis, KIIs
National Girls' Education Communication Strategy 2014+	GESI Analysis
National Education Sector Plan 2008–2018 and the Education Sector Implementation Plans 2013/2014–2017/2018+	GESI Analysis, KIIs
Re-admission Policy+	GESI Analysis, KIIs
School Child Protection Policies+	GESI Analysis
National Inclusive Education Strategy 2016–2020+	GESI Analysis, KIIs
MOEST Education Sector Implementation Plan II+	GESI Analysis
Adolescent Girls and Young Women Strategy+	GESI Analysis
National Gender Policy 2015+	GESI Analysis
Gender Equality Act 2013+	GESI Analysis
Education for All Policy	KIIs
Free Primary Education Policy	KIIs
School Child Protection Policy	KIIs
Orphan and Vulnerable Children Policy	KIIs
Social Protection Policy	KIIs
Child Protection Framework	KIIs
National Gender Policy	KIIs
Gender Equality Act	KIIs
National Reading Programme	KIIs
National Education Standards	KIIs
Sexual and Reproductive Health Rights	KIIs
Gender-Based Violence Act	KIIs
Adolescent and Youth Women Strategy	KIIs
Disability Act	KIIs

Note: A plus sign (+) indicates that the policy was included in the baseline indicator calculation.

Regarding implementation of policies, one district-level and one national-level official noted that there is no uniform approach to supporting girls, explaining that each school decides how best to support girls through Girls' Clubs or similar activities. A national-level official said that adolescent girls have access to CBE regardless of whether they have children or not and that young mothers can bring children to be cared for by Learning Centre Management Committee members.

Generally, government officials expressed significant concerns over lack of resources to implement plans and policies to support marginalised girls' education. Specific impediments mentioned included including understaffed offices, logistical challenges, capacity gaps, lack of technology and limited monitoring resources.

IO5.2 Number of District Education Conferences which reference marginalised girls' education

According to KII respondents, there is one district education conference per year, primarily organised by nongovernmental partners. Respondents mentioned that dissemination of information, such as minutes or resources, often are at the discretion of the organiser. National-level officials were not aware of any district education conferences.

IO5.3 National and district government staff demonstrate supportive attitudes to marginalised girls' education

Baseline attitudes regarding marginalised girls' education were explored through KIIs. Respondents were asked about their level of interest in strengthening the support and programming for marginalised girls. Although district officials expressed a high level of interest, they caveated that they could not speak for the national level. One district official said that the level of interest and awareness of programming varied across ministries. Specifically, the respondent said that technical staff tend to be more aware and interested than support staff. A national-level official said that staff at the national level are very interested in supporting marginalised girls' education and are well informed of the policies and mechanisms for support.

When asked about strategies for getting ministry staff interested involved in marginalised girls' education, examples included presentations and CBE site visits, trainings and capacity-building activities and awareness and sensitisation campaigns.

Reflections

Baseline results for IO5 indicators were primary explored qualitatively to ensure if they are appropriate measurements of TEAM Girl Malawi success under this IO. Data from the GESI was also incorporated, specifically for IO5.1. Given the small sample of government officials interviewed, and due to their wide range of knowledge about marginalised girls' education, it was difficult to fully understand the present levels of district and national leadership and engagement. Minutes of district education conferences were not available and may be difficult to obtain for future evaluation points. Further, staff attitudes related to IO5.3 were explored in KIIs but were not quantified using a formal survey or questionnaire due to budgetary limitations at baseline and a limited number of potential respondents.

It is unclear how much the indicators as stated will be able to capture strengthening of government support in marginalised girls' education. The respondents identified by the project did not appear to have a strong understanding of the topic or how institutional change could be made. TEAM Girl Malawi should evaluate their strategy for strengthening government support and ensure that it has the potential to lead to the changes being measured in their selected indicators.

Specifically, STS recommends the following updates to indicators under IO4:

- IO5.3: Percentage of national and district government staff that indicate supportive attitudes to marginalised girls' education

At the next evaluation point, the following actions should be taken to ensure that indicator data is adequately collected:

- TEAM Girl Malawi should define what policies and plans are considered relevant for IO5.1 and stay updated on any new policies or plans that are released over the life of the project
- TEAM Girl Malawi should aggregate district education conference meeting minutes over the life of the project to best respond to IO4.2

- TEAM Girl Malawi should identify the key government stakeholders who will be involved in the project in coming years, so these respondents can be targeted for data collection at the next evaluation point

At the next evaluation point, new tools or items will be introduced to track indicators under IO4. These include:

For IO5.3:

- Government official survey

Qualitative data for IO5 indicators will be captured from district and national-level government officials. Efforts should be made to target interviews to officials with the greatest interaction and knowledge of TEAM Girl Malawi and marginalised girls' education initiatives. Efforts will also be made to have STS conduct KIIs with government officials at coming evaluation points so that data gathered is useful and comprehensive.

Targets

Based on the data and evidence collected at baseline, recommended targets for IO5 indicators are:

- IO5.1 Number of national & district plans & policies which include measures to strengthen primary schools to meet the needs of marginalised adolescent girls: 1 additional policy
- IO5.2 Number of District Education Conferences which reference marginalised girls' education: 2
- IO5.3 Percentage of national and district government staff that indicate supportive attitudes to marginalised girls' education: 75%

Project Checks on Intermediate Outcomes

Ensure that the IO analysis reflects the links between different levels in the logframe and informs the validity of the Theory of Change. This includes checking whether the EE (?) have:

- Measured and analysed all IO indicators presented in logframe.
- Disaggregated the data according to the logframe.
- Used both the qualitative and quantitative analysis stated in the logframe.
- Related the IO analysis to the analysis of Outcomes.

Project confirmation that all the IO analysis has been covered.

7.2 Life skills

TEAM Girl Malawi indicator O1.3—number of highly marginalised girls supported by GEC with improved life skills outcomes—was measured by creating a composite index. The index was comprised of domains specifically related to the TEAM Girl Malawi Girls' Clubs curriculum. It also builds on IOs lower in the project's ToC.

Specifically, the life skills index contained items from the following domains already measured and reported under the IOs: attitudes towards education, self-esteem, self-confidence, CP knowledge and attitudes, attitudes towards gender-based violence and SRHR knowledge,

attitudes and practices (see Supplemental Table 36 for list of items). A total of 361 girls provided responses to the items on the survey.

To calculate baseline levels of life skills, each girl's mean score on the life skills index was computed on a 3.00-point scale. Then, girls' scores were categorised as high and low. High life skills scores were defined as scores greater than 1.75, the median of the life skills index.

Main findings

Findings for O1.3 are presented in Supplemental Table 31. The overall mean score for girls on the life skills index was 1.71 out of 3.00. Nearly half (49.86%) of girls had a high life skills score. Proportions of girls receiving high scores varied across districts. Specifically, 60.00% of girls in Dedza received a high score, while 33.02% of girls in Mchinji received a high score. Age also appeared to be important. The proportion of girls receiving high scores increased as the age range increased—24.56% for girls aged 10 or 11, 44.88% for girls aged 12 through 16 and 74.75% for girls aged 17 through 19. Additionally, 37.04% of girls who experienced bullying had a high score.

Scores on the life skills index are strongly correlated to IO2.1 and IO2.2. Life skills scores are moderately correlated with IO4.3 (Supplemental Table 44). This is because the index for O1.3 uses many of the same items for indices for IOs 2.1, 2.2 and 4.3. As a result, the index may not provide a different understanding of girls' life skills beyond the results in the IOs.

Supplemental Table 31: O1.3 Results by subgroup and barrier (median of 1.75 out of 3.00)

Category	N	Score	Proportion of total
All girls	**361**	**Low score**	**50.14%**
		High score	**49.86%**
Lilongwe	75	Low score	40.00%
		High score	60.00%
Dedza	180	Low score	50.67%
		High score	49.33%
Mchinji	106	Low score	66.98%
		High score	33.02%
Age 10–11	57	Low score	75.44%
		High score	24.56%
Age 12–16	205	Low score	55.12%
		High score	44.88%
Age 17–19	99	Low score	25.25%
		High score	74.75%
Subgroup			
Married and caregiver	70	Low score	21.43%
		High score	78.57%
Orphaned	95	Low score	45.26%
		High score	54.74%
Head of household	11	Low score	54.55%
		High score	45.45%
High chore burden	174	Low score	47.13%
		High score	52.87%
Girls with disabilities	125	Low score	51.20%
		High score	48.80%
Barrier			
Bullying	27	Low score	62.96%

Category	N	Score	Proportion of total
		High score	37.04%
School cost	329	Low score	50.46%
		High score	49.54%
Parent support	67	Low score	50.75%
		High score	49.25%
Menstruation	108	Low score	43.52%
		High score	56.48%
Food insecurity or hunger	234	Low score	49.57%
		High score	50.43%
School safety	102	Low score	53.92%
		High score	46.08%

STS ran a regression model to understand the relative predictive influence of a set of factors on girls' life skills scores. Results are presented in Supplemental Table 32. These factors included age, district, hunger and married and caregiver status (see Section 4.4 Quantitative evaluation methodology for explanation on inclusion and exclusion of variables).[116] Findings indicate that having a girl in the 12 through 16 or 17 through 19 age groups was a statistically significant predictor of girls' life skills, as was being from Mchinji and being married and a caregiver, controlling for experiencing food insecurity or hunger as a barrier. If a girl is 12 through 16 years old, her life skills score is likely to be 0.28 points higher than girls aged 10 or 11. Similarly, if a girl is 17 through 19 years old, she is likely to score 0.39 points higher than a girl who is 10 or 11 years old. Girls who are married and caregivers are likely to have scores that are 0.21 points higher than girls who are not, controlling for experiencing food insecurity or hunger as a barrier. Girls from Mchinji are likely to score 0.19 points lower than girls from all other districts. Hunger was not a statistically significant predictor of girls' life skills scores.

Supplemental Table 32: O1.3 Girls' life skills analytical model results

Category	Coefficient	Standard error	95% Confidence interval	
			Min.	Max.
Age 12–16	0.28***	0.06	0.16	0.40
Age 17–19	0.39***	0.07	0.24	0.53
Mchinji	-0.19***	0.05	-0.28	-0.09
Married and caregiver	0.21***	0.06	0.09	0.33
Food insecurity or hunger	-0.03	0.05	-0.12	0.05
Constant[117]	1.48	0.06	1.35	1.60

Note: One asterisk (*) denotes differences between groups that are statistically significant at $p < 0.05$. Two asterisks (**) denote differences that are statistically significant at $p < 0.01$ and 3 asterisks (***) denote differences that are statistically significant at $p < 0.001$.

Qualitative findings related to girls' life skills are explored in Section 7 under IO1, IO2, IO3 and IO4.

Reflections

[116] F (5, 355) =16.93, p<0.001

[117] The constant, or intercept, is the average score for the reference group. In this case, the reference group is girls who are 10 or 11 years old from Dedza or Lilongwe who are not married caregivers and whose households have not gone to bed hungry 10 or more days in the last year.

The life skills that TEAM Girl Malawi intends to strengthen through TfaC's Girls' Club sessions are:

- Communication (year 1)
- Reproductive health (year 1)
- Sexual health (year 1)
- Child rights (year 2)
- Gender rights (year 2)
- Planning for the future (year 2)

These skills are captured through IOs in the TEAM Girl Malawi logframe. The selection of these topics was driven by in-depth knowledge by implementers of adolescent girls' needs in target communities, as well as by the needs' assessment conducted by TEAM Girl Malawi. Baseline data collection was targeted to measure these areas, particularly through IO indicators. Given the findings, these appear to be life skills that girls are presently lacking and that are related to learning outcomes (see Supplemental Table 7). The life skills interventions proposed by the project are supportive of improving girls' agency and rights. Based on the information available, the proposed life skills curriculum appears to be GESI responsive. However, it is not evident that measuring life skills as an outcome through indicator O1.3 provides information above and beyond what is measured through other IO indicators. TEAM Girl Malawi should assess whether there are additional domains of life skills that should be measured beyond what is in IOs.

At the next evaluation point, new tools or items will be introduced to track improvements in life skills, including:

- Self-confidence observation checklist (developed and piloted by TfaC)
- Additional girls survey items

Additional questions will be added to FGDs with adolescent girls, and questions will be asked of CBE facilitators to understand qualitatively the explanatory factors around girls' life skills improvements over time.

Project response

We have confidence in the index composition, but there is the potential for over estimation in this area when relying on self-reported data. We will triangulate these findings with internal monitoring data – TfaC are using a mix of participatory and observational tools to measure levels of SRHR knowledge, attitude and skills as well as self-confidence and self-esteem; the results from this process will provide a fuller understanding of the girls life skill levels, particularly regarding SRHR and self-confidence. Further analysis and programme adaptation will ensue in the coming weeks/months, and the Girls' Club curriculum is designed to tackle these pertinent issues, and we will be monitoring its impact on these key areas of learning.

8. Conclusions

This baseline report presents comprehensive, mixed-method evidence on the current status of outcomes and IOs for TEAM Girl Malawi cohort 1 beneficiaries. A summary of the findings and implications for the planned interventions are detailed.

Key characteristic subgroups and barriers

TEAM Girl Malawi conducted a needs assessment prior to enrolling beneficiaries to identify the subgroups that are marginalised within their target communities. The enrolment process conducted by the project effectively targeted girls within these all of these subgroups, except for girls who have albinism. Girls who are heads of household comprised a small proportion of the sample (2.91%), although nearly half (46.30%) of girls in the sample identified as the primary caregivers for her own or other children.

The barriers to learning and transition analysed throughout the baseline report were identified using a mixed-methods approach, in which girls' qualitative data was utilised to highlight barriers. Following the identification of these barriers, qualitative responses were used to calculate the proportions of girls experiencing these barriers (see Annex 15: Additional tables Supplemental Table 35 for list of items included in each barrier). Two of the most frequently experienced barriers—school costs and food insecurity or hunger—are linked to households' levels of poverty. Menstruation was also a frequently experienced barrier; 33.07% of girls experienced menstruation as a barrier. School safety was cited by 27.25% of girls, and lack of parental support for school was cited by 18.52% of girls.

TEAM Girl Malawi programming appears to be highly sensitive and proactive to responding to these barriers. SRHR, safety and community engagement in girls' education are all key themes of the project's ToC. Additional project inputs related to mitigate financial barriers—such savings groups' engagement and microloan disbursements— also target beneficiaries.

Learning outcomes

Girls' baseline literacy levels are notably low. Only on 2 subtasks—letter name identification and listening comprehension—did a majority of girls answer at least one item correctly. On all other subtasks—initial sound identification, syllable identification, familiar word reading, oral reading fluency and reading comprehension—more than 40% of girls did not answer a single item correctly. Results on listening comprehension—on which 44.71% of girls scored as established learners and 20.37% as proficient learners—indicate that TEAM Girl Malawi beneficiaries have a command of oral vocabulary and oral language comprehension. Their comprehension difficulties are a result of being non-readers.

Girls' numeracy scores at baseline were more encouraging. About a quarter (26.19%) of girls scored as proficient learners on number recognition, and more than one-third of girls scored as established learners on the number recognition, quantity discrimination, addition level 1 and word problems subtasks. Girls appeared to have more challenges with number patterns—measured by the missing number subtask—and higher order arithmetic—measured by addition level 2 and subtraction level 2.

Learning outcomes by subgroups and barriers indicated that girls aged 12–16 and girls aged 17–19 both outperformed girls aged 10–11. Girls who had functional difficulties walking, girls who had functional difficulties communicating or girls who reported bullying as a barrier all had statistically significantly lower literacy and numeracy aggregate scores than did all other girls. Girls in the

married and caregiver, orphaned or high chore burden subgroups had statistically higher literacy and numeracy scores than did all other girls. There were no significant differences in literacy and numeracy scores by district.It is unclear how the project will target learning sessions to the different literacy and numeracy levels or subgroups of learners, though given the distribution in performance, the project should consider taking a differentiated approach, particularly to mathematics teaching.

Transition outcomes

Most of the girls in cohort 1 (84.55%) attended some school before enrolling in TEAM Girl Malawi. A majority (61.38%) of girls in the sample had reached standard 4 or 5. Larger proportions of younger girls in the sample—namely, those from age group 10–11 and transition group A—had never been to school before enrolling in the project. Notably, 37.60% of girls in the sample who experienced food insecurity or hunger as a barrier had never been to school. Nearly all girls in the baseline sample (94.40%) said they believed they would finish CBE, and more than one-half hoped to go to vocational training after finishing. A smaller proportion of girls—about one in 5—expressed a hope to return to primary school after completing CBE. Overall, larger proportions of girls in younger age groups reported that they hoped to return to primary school after competing CBE—29.63% of girls aged 10–11—and larger proportions of girls aged 17–19 groups reported that they hope to go to vocational training or work in safe, fairly paid jobs—58.75% and 42.27%, respectively.

These findings indicate that girls of different ages have different expectations for transitions after CBE, which aligns with TEAM Girl Malawi transition pathways. Given that only about 1 of 3 girls aged 10-11 hope to return to primary school, the project should focus attention on changing perceptions of the opportunities gained through the formal school system.

When evaluating the project's planned transition pathways through a GESI lens, the pathways appear to be gender accommodating. Given current evidence, it is not clear to what extent the project will create enabling conditions for transition for girls with social differences or inequalities, particularly for those with disabilities. The project should ensure that it is encouraging inclusive education opportunities in the formal school system for girls with disabilities who choose to continue through primary school. Further, TEAM Girl Malawi should encourage vocational training and employment opportunities that are accessible to all girls, regardless of their functional difficulties. Further, the project may consider how to mitigate the barriers faced by young mothers, who may be restricted from engaging in formal education and training opportunities due to their childcare responsibilities.

Sustainability outcomes

Sustainability findings at baseline—presented for system, community and learning space indicators—were drawn primarily from qualitative data. The overall score on the sustainability scorecard was 1.00 out of 4.00. Overall, evidence suggested mixed levels of enabling environments for sustainability. System-level sustainability evaluated district- and national-level education officials' knowledge about and responsiveness to marginalised girls' educational needs. Findings varied across respondents, but all respondents named lack of resources as an impediment to supporting and sustaining initiatives for marginalised girls' education. Evidence on community sustainability was more encouraging. Community leaders and other stakeholders exhibited knowledge of and willingness to engage in sustaining marginalised girls' education. Learning-space evidence at baseline was limited. Further research into these indicators will be collected and analysed at the next evaluation point.

Intermediate outcome findings

IO findings at baseline provide substantial data that can be used to formatively target TEAM Girl Malawi interventions. Although indicators under IO1 and IO3 were 0.00, as learning sessions had not yet begun, qualitative findings highlighted a number of considerations that TEAM Girl Malawi should take into account to ensure regular attendance and effective learning. Girls' SRHR understanding was notably low at baseline, while their self-esteem and self-confidence were relatively higher. Findings related to community support showed moderate levels of existing support, with room for growth over time.

Gender equality

TEAM Girl Malawi interventions fulfil the requirements of 'gender sensitive'. The project conducted a targeted, inclusive enrolment process that reached the intended subgroups of marginalised girls. It addresses many of the identified barriers that restrict girls' learning and transition, and it is aimed at changing perceptions throughout the communities where it is working. By including equally marginalised boys in CBEs, the project is aiming to reduce resentments and perceptions of favouritism while also allowing support for social-norm change and equality. The project should remain cognisant of gender dynamics in the classroom, especially as there will be a range of ages of girls and boys in the learning environment. Girls and boys should also be actively engaged together in Girls Club sessions, to ensure that sensitive topics can be explored in a gender transformative ecosystem.

Theory of change

Assumptions in the ToC regarding subgroups and barriers appear to hold true. The most prevalent social, economic and educational barriers uncovered through the baseline are considered in TEAM Girl Malawi intervention planning. These include support for girls' SRHR—specifically menstrual health—through Girls' Clubs, financial support through micro-loans for households with poverty or food insecurity and system-level support for families.

Not all the barriers listed in the project's ToC were evidenced through the baseline data—namely, the educational marginalisation barriers. This is primarily due to the respondents that participated in the baseline – specifically, out-of-school girls and their family members instead of stakeholders and beneficiaries in the formal school system. Nevertheless, findings from the baseline do not explicitly disprove any of the educational barriers that were identified by the project through the GESI, and several were substantiated through qualitative findings. TEAM Girl Malawi may want to revisit assumed educational barriers through monitoring to ensure that they continue to be applicable to the beneficiary population and communities.

Risks

Given the high level of sensitivity of TEAM Girl Malawi beneficiary girls, the project should be aware of any heightened stigma or security threats that arise for the girls who are attending CBEs. Girls and their caregivers noted safety and security at and on the way to school as barriers, so the project should closely monitor any threats faced by participants as a result of their attendance. Given mentions of physical and sexual violence against girls, the project should also ensure proper safeguarding training, particularly of staff, to be aware of signs and reporting mechanisms. Further, TEAM Girl Malawi should monitor gender power balances in CBE classrooms. Although girls will outnumber boys in learning sessions, cultural norms indicate that boys show greater confidence and are more outspoken. CBE facilitators should use gender-sensitive pedagogy and

ensure that all learners in CBE classrooms are engaged and contributing. Similar risks and mitigation efforts related to gender balance should be taken for Girls Club.

9. Recommendations

This section provides recommendations to TEAM Girl Malawi and reflections for the evaluation of the project resulting from baseline findings.

Monitoring, evaluation and learning

- A key recommendation for TEAM Girl Malawi is to establish streamlined, routine and rapid systems for monitoring beneficiary attendance and progress through CBE and into transition pathways. TEAM Girl Malawi's enrolment data and high response rates at baseline indicate that, to-date, the project has succeeded in enrolling and engaging its target population. However, it will be critical to set up equally strong monitoring systems to measure beneficiary attendance and progress early and often. Further, to allow for the longitudinal evaluation design to measure girls through CBE and transition pathways, the project should institutionalise the use of girls' unique project identifiers—perhaps with cards or badges.
- Results for IO2 indicated that girls' SRHR understanding at baseline was notably low, while their self-esteem and self-confidence were relatively higher. However, given the qualitative findings that indicate girls feel a lack of bodily autonomy, it is unclear if these results adequately measure the intended constructs. Self-esteem and self-confidence results should be explored through routine monitoring of girls' behaviours, as these incongruent results may have been driven by the specific items included in surveys. TfaC observational data (Annex 16: TfaC baseline results summary) indicated that SRHR and self-confidence may be lower than reported directly by girls. These findings should be utilised to improve upon surveys at the next evaluation point and ensure that IO2 indicators are measured as accurately as possible.
- Indices for the IO indicators were created by mapping survey items to indicators. Reliability of the indices, using Cronbach's alpha, indicated that least 2 of the indices had very low alpha values—IO4.2 SRHR support and IO4.3 girls' education support. This means that that either too few items comprised the composite, items were not clearly understood by respondents or composite items inadequately measured the underlying construct. While these composites were constructed from survey items—and alternative items were not always available—revised items, additional items or additional responses may be necessary to improve reliability at midline.
- TEAM Girl Malawi's current evaluation questions are comprehensive. The project may consider examining evaluation questions to ensure that they best fit the learning agenda of the intervention. This would be beneficial if there are specific components of interest in the project's intervention.
- The project should evaluate IO5 to ensure that indicators best capture the intended and expected outcomes of TEAM Girl Malawi's activities aimed at district and national government stakeholders. Recommended modifications or replacements to indicators are suggested in Section 7.
- Outcome 1.3—number of highly marginalised girls supported by GEC with improved life skills outcomes—was measured by creating a composite index comprised of domains already measured and reported under the IOs. As a result, it is not evident that measuring life skills as an independent indicator provides information beyond what is measured through other IO indicators. TEAM Girl Malawi should assess whether there are additional domains of life skills that should be measured beyond what is being measured through IOs based on the project's planned activities.

Design

- Disability results from the Washington Group Short Set of Disability Questions—administered during enrolment—and from the Washington Group/UNICEF Module on Child Functioning—used at baseline—led to notably different prevalence. Specifically, the proportion of girls with at least one domain of functional difficulty was 8.47% at enrolment but 34.39% at baseline. While some of this difference may be due to the increased number of domains considered on the Child Functioning module, results from screenings conducted by TEAM Girl Malawi after enrolment only identified a small number of girls as having disabilities.[118] The project should train CBE facilitators to closely monitor learners' and identify potential disabilities at the outset of CBE learning sessions. This will better ensure that learners are screened for disabilities in a targeted way and that learners are provided with the supports they need to progress through the CBE programme.
- Chichewa literacy levels of project beneficiaries are very low, and numeracy were only slightly better. Because CBE centres will be serving learners across a wide range of ages and baseline literacy and numeracy levels, the project should focus on training CBE facilitators in differentiated approaches to ensure that all beneficiaries progress according to their specific needs and skill levels.
- Although girls' expectations for pathways after completing CBE generally align with the transition pathways of TEAM Girl Malawi based on their age, only 20.69% of girls at baseline said they hoped to go to primary school. Larger proportions of girls aged 10–15 (25.87%) and of girls from Dedza (31.51%) said they hoped to go to primary school, the project should focus on increasing the educational aspirations of girls aged 10–15 and the support of their caregivers for education.
- Girls' understanding of SRHR was notably low at baseline—mean of 4.06 out of 18 points. Additionally, a regression analysis indicated that girls aged 10–11 were more likely to have lower SRHR understanding. Qualitative data from girls, however, indicated that girls are believed to experience their sexual debut starting at age 12. The project may consider tailoring SRHR curriculum for younger girls to ensure that they gain critical knowledge before they begin to engage in sexual activity.
- Pre-existing ideas about learning environment and gender may impact CBE learning sessions. Learners expressed concerns about relationships between facilitators and teachers with learners. Concerns about animosity or poor relationships were frequently mentioned during qualitative research, as were fears of corporal punishment. Further, qualitative findings from girls and boys emphasised that adolescents view peers of the opposite sex as receiving different treatment—whether positive or negative—and that boys and girls act differently when around each other. TEAM Girl Malawi staff should be sensitised and well-trained to address these challenges and to negate any feelings of favouritism.
- Safety on the way to and at CBE is a concern for beneficiaries and caregivers. The girls' survey asked if respondents felt safe travelling to and from school, and 46.15% reported that they did not. Of the 353 caregivers responding to the household survey, 24.08% reported that it was fairly or very unsafe for girls to travel to schools in their area. Qualitative data highlighted that learners anticipated bullying, fighting and physical violence at schools. The project should consistently monitor safety concerns throughout the life of the project.

Sustainability

[118] Only nine girls (2.4%) at baseline required large print due to a visual disability identified during the screening.

- Initial findings on systems-level sustainability indicated that the project may face challenges in engaging district and national leadership, while community-level sustainability findings highlighted those stakeholders' knowledge of and willingness to engage in sustaining the educational opportunities of marginalised girls. The project should evaluate and tailor its approach to engaging local government and local stakeholders in a way that will ensure their buy-in of TEAM Girl Malawi implementation. For instance, they may engage education officials in data collections and trainings.