coffey
A TETRA TECH COMPANY

Endline Evaluation Report

Step Change Window

Final Report (December 2017)

Evaluation Manager Girls' Education Challenge (GEC) Fund

ORB International RTI INTERNATIONAL

Endline Evaluation Report

Step Change Window

Final Report (December 2017)

Evaluation Manager Girls' Education Challenge (GEC) Fund

Lead Authors

- Simon Griffiths
- Florian Poli
- Sophie Amili

Contributions by

- Paul Diegert
- Panayiota Kastritis
- Karolin Krause
- Brian Law
- Ella Page

This document has been approved for submission by Coffey's Project Director, based on a review of satisfactory adherence to our policies on:

- Quality management
- HSSE and risk management
- Financial management and Value for Money (VfM)
- Personnel recruitment and management
- Performance Management and Monitoring and Evaluation (M&E)

Ben Ward, Project Director

Signature:

Disclaimer

This report is provided on the basis that it is for the use of the Department for International Development only. Coffey International Development Ltd will not be bound to discuss, explain or reply to queries raised by any agency other than the intended recipients of this report. Coffey International Development Ltd disclaims all liability to any third party who may place reliance on this report and therefore does not assume responsibility for any loss or damage suffered by any such third party in reliance thereon.

Abbreviations and Acronyms

ABE	Alternative Basic Education
ACTED	Agency for Technical Cooperation and Development
AIDS	Acquired Immune Deficiency Syndrome
AKF	Aga Khan Foundation
ALP	Accelerated Learning Programme
ASER	Annual Status of Education Report
BEAM	Basic Education Assistance
BEEP	Bicycle Education Empowerment Programme
BL Report	Baseline Report
CBE	Community-Based Education
CfBT	Education Development Trust
DAC	Development Assistance Committee
DFID	Department for International Development (United Kingdom)
DIBELS	Dynamic Indicators of Basic Early Literacy Skills
DID	Difference-in-Difference
DRC	Democratic Republic of Congo
EGMA	Early Grade Math Assessment
EGRA	Early Grade Reading Assessment
EM	Evaluation Manager
FGM/C	Female Genital Mutilation/Cutting
FM	Fund Manager
GEC	Girls' Education Challenge Fund
GPE	Global Partnership for Education
HHS	Household Survey
HIV	Human Immunodeficiency Virus
IDI	In-depth Interview
IDPs	Internally-Displaced Persons
IRC	International Rescue Committee

IW	Innovation Window
MDGs	Millennium Development Goals
M&E	Monitoring and Evaluation
ODA	Overseas Development Aid
ORB	Opinion Research Business
OOS	Out-Of-School
PbR	Payment-by-Results
Plan	Plan International
RCT	Randomised Controlled Trial
RI	Relief International
RTI	RTI International
SBA	School-based assessment
SCW	Step Change Window
SD	Standard Deviation
SPW	Strategic Partnerships Window
STAGES	Steps Towards Afghan Girls Educational Success
STC	Save the Children
SVS	School-visit survey
ToC	Theory of Change
UN	United Nations
UNICEF	United Nations International Children's Emergency Fund
UNESCO	United Nations Educational, Scientific and Cultural Organisation
UNGEI	United Nations Girls' Education Initiative
US	United States of America
USAID	United States Agency for International Development
VfM	Value for Money
WPM	Words per Minute
WUSC	World University Service of Canada
WV	World Vision

Table of Contents

Executive Summary

Introduction to the GEC

The Department for International Development (DFID) launched the £355 million Girls' Education Challenge Fund (GEC) in 2012, setting an ambitious target of reaching one million marginalised girls by March 2017[1].

The Business Case[2] for the GEC recognised that there was a lack of robust evidence of the causes of girls' marginalisation from education. This was the underlying rationale for a programme that not only funded interventions, but challenged organisations to identify the causes of educational marginalisation and rigorously collect evidence of projects' performance to ascertain what worked well, why, for whom and under what circumstances.

DFID appointed Coffey (in partnership with RTI International and ORB), as the Evaluation Manager (EM), with responsibility for independently evaluating the overall effectiveness and impact of the programme. The EM has closely collaborated with the GEC Fund Manager (FM) in supporting projects to collect data and report their results consistently across the programme.

In July 2016, DFID announced it would provide a further £100 million of funding to the GEC[3] (through the GEC Transitions Window) to continue supporting one million marginalised girls to transition through school phases and complete a cycle of education, and to support an additional 175,000 of the poorest and most marginalised girls to receive a quality education (through the Leave No Girl Behind Window)[4].

The Step Change Window

The Step Change Window (SCW) is one of three funding windows in the GEC, the others being the Innovation Window (IW) and Strategic Partnerships Window (SPW).

SCW projects were awarded funding of up to £30 million per project to apply tried and tested designs that could quickly and effectively expand education opportunities and improve the quality of education for girls at primary and secondary school levels. They operate in nine countries: Afghanistan; DRC; Ethiopia; Kenya; Mozambique; Sierra Leone; Somalia; Tanzania; and Zimbabwe. Fourteen SCW projects were operating at the time of the endline evaluation, aiming to address multiple barriers to girls' education affecting individual girls, their households, communities, and schools.

Purpose of the endline evaluation report

The purpose of this endline evaluation report is to provide reliable evidence of the programme's impact on girls' access to education and their learning outcomes, as well as projects' effectiveness, and the sustainability and value for money (VfM) of their interventions. These findings will be used by DFID, the FM and projects to further develop projects' activities funded by the GEC, in particular for the additional eight years planned beyond the end of this phase[5]. This report follows the SCW Baseline Report published in January 2015 and the SCW Midline Report published in March 2017, and covers the first three-year phase of the GEC implementation period. This report should also generate transferable learning for a wider audience, including donor agencies, governments of GEC target countries, and other policy-makers.

---

[1] Girls' Education Challenge: http://devtracker.dfid.gov.uk/projects/GB-1-202372
[2] GEC Business Case v.3, September 2011
[3] For more details see DFID's press release on the extension of the GEC, at: https://www.gov.uk/government/news/britain-to-help-175000-girls-in-worlds-poorest-countries-get-an-education.
[4] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/559123/leave-no-girl-behind.pdf.
[5] GEC Business Case v.3, September 2011

Evaluation and research approach

The GEC Evaluation Strategy, produced by the EM in late 2012, set out a two-pronged approach to evaluating the SCW. The EM's research focuses on the effectiveness and impact of activities delivered by all projects across the window on the catchment communities they targeted. A random sample of girls was selected, who live in the catchment areas of the treatment and control schools for all fourteen SCW projects and were aged 5-15 at baseline (8-18 at endline). The samples are designed to be representative of the general SCW target population, i.e. large enough to detect and measure the impact of the projects on their communities at the window level. They are not large enough to reliably measure the impacts of each individual project. For this purpose, the FM and EM supported projects and their externally commissioned evaluators to design their own separate research samples that would be large enough to measure impact on their specific target groups. Furthermore, many of the projects designed their samples to purposefully test the girls that were participating in project activities. When interpreting and presenting findings from the EM data that relate to individual projects, we have been careful to stress that these only provide an indication of the potential effects at this level.

The EM's primary research for the SCW endline evaluation consisted of: 6,279 longitudinal household surveys, literacy (EGRA[6]) and numeracy (EGMA[7]) tests; and 3,107 school visit surveys in girls' schools. The EM also conducted reviews and meta-synthesis of the endline evaluation reports, datasets and outcome spreadsheets submitted by SCW projects to triangulate the analysis of the EM's quantitative data.

Key findings

Reach and equity

Most SCW projects reached their target number of beneficiaries at endline[8]. While we cannot ascertain whether SCW projects targeted the most marginalised girls in their communities or regions, our data shows that girls are disadvantaged across a variety of educational, social and economic factors whose prevalence and severity differ across project areas. This reflects DFID's approach to allowing projects themselves to define how their target girls were marginalised from education according to the contexts in which they lived, which is consistent with the findings from the GEC Process Review.[9] Although some projects targeted specific subgroups of marginalised girls, the extent to which projects reached those subgroups and improved their learning remains unclear. In fact, with the exception of out-of-school girls, very few of the subgroups were tracked as a specific cohort. Activities aiming to improve the learning levels of different subgroups of girls have to be context-specific and tailored to girls' varying needs. This highlights the importance of analysing contextual factors as an integral part of project design and M&E processes – in particular, how and why different factors impact on the education of different subgroups of marginalised girls, and the extent to which they constrain their capacity to stay in school and learn effectively.

Impact on learning outcomes

The FM reports[10] that the SCW 'substantially met' its learning outcome target by reaching and benefiting 577,618 girls with improved learning outcomes[11], an 88 percent achievement against the endline target. While the SCW has reached and benefited a substantial number of marginalised girls, several projects did not have enough of an effect on girls' learning to achieve their targets. SCW learning targets were typically defined as an improvement of 0.5 SD over the three-year implementation period of the GEC[12]. Eight out of fourteen SCW projects met their literacy targets between baseline and endline, while only four projects met their numeracy targets.

The EM data shows that literacy and numeracy levels steadily improved among girls in SCW target communities. On average, girls' oral reading fluency improved by 16 words-per-minute (EGRA wpm) and their numeracy

---

improved by 25 points (EGMA/100) since baseline. However, similar increases are observed in control communities. These findings hide disparities across projects and between subgroups. In some targeted communities, large proportions of girls still cannot read a single word-per-minute in their language of instruction at endline, while in other communities most girls can read fluently. Overall, our EM data shows that lower-performing girls have improved more than other girls of the same age, suggesting that SCW projects were successful in targeting and reaching the most educationally marginalised. Similarly, the learning levels of out-of-school girls have improved more than girls who are in school.

Despite these improvements, the EM data shows that half of the girls in the communities targeted by the SCW achieved a reading fluency level of 45 words per minute by their sixth year of primary education or later. International benchmarks for English as the language of instruction suggest that children need to achieve this by their second year of primary school. Reading fluency is essential for comprehension, and many girls are spending most of their primary school years learning to read rather than reading to learn. This not only constrains girls' literacy but their ability to progress and transition through school grades with the functional skills needed to learn effectively. Although improvements in literacy and numeracy scores seem largely correlated, SCW projects found it harder to improve girls' numeracy, as demonstrated by the small number of projects meeting their numeracy targets. This may be explained by a lack of teachers with the specific "hard" skills needed to teach numeracy that may not necessarily be required to teach literacy[13].

Impact on attendance

The EM data shows that in-school attendance rates have remained constant overall across SCW target communities since baseline[14]. Enrolment rates increased slightly and to a similar extent in treatment and control areas, although some evidence suggests that projects have been able to slow down drop-out rates of older girls. However, this window-level community data masks significant differences at the project level, in particular that the proportion of out-of-school girls among target communities is much higher in some projects' contexts (Somalia, Ethiopian Afar region, Kenyan refugee camps). Further analysis shows that in these contexts and in Afghanistan, most out-of-school girls have never attended schools, while in other projects' contexts, out-of-school girls are older girls who dropped out after completing primary school.

The difficulty of capturing small changes through household-based measures has meant we do not have conclusive findings at endline about improvements in girls' attendance at school across the SCW. The FM also reported inconsistencies and poor quality of the school-based measures of attendance used by projects, largely due to the unreliability of school records and registers, and restrictions to spot checks in some schools. We acknowledge this and faced similar challenges during our research in schools.

Effectiveness of SCW projects' activities

The effects of particular interventions on girls' attendance and learning was often inconclusive, due to a lack of sufficient and reliable evidence at the project level linking individual activities and outputs with specific effects on learning outcomes. This is partly due to the holistic approach taken by projects, which made it very difficult to quantitatively identify the effects of specific activities. Based on the available evidence, we found that economic interventions (i.e. scholarships, bursaries and stipends) and provision of in-kind support helped to reduce the cost of schooling and led to improved attendance among girls. However, the amount of time spent by girls on household chores does not seem to have reduced as a result of projects' interventions, and domestic responsibilities still compete with the time girls spend in school.

Activities aimed at directly improving learning seem to have had the biggest impact on literacy and numeracy. Infrastructure and facilities improvements positively impacted on attendance, and the provision of learning materials improved learning. Community-based activities (such as gatherings and meetings) have encouraged caregivers and community members to take an active role in supporting girls' education, leading to improvements

---

[13] Atweh et al (2014) highlight that for early numeracy teaching, teachers must develop connected understandings of early numeracy content and problem solving, knowledge of the progression of numeracy learning development, syntactic knowledge related to problem solving as well as pedagogic content knowledge and skills related to early numeracy teaching. See *Bill Atweh, Arindam Bose, Mellony Graven, Jayasree Subramanian, Hamsa Venkat (2014). Teaching Numeracy in Pre-School and Early Grades in Low-Income Countries. GIZ.*

[14] Enrolment rates slightly increased, but they did so to a similar extent across treatment and control areas. But enrolment, as opposed to attendance, is not a core GEC outcome.

in school attendance, but it is unclear whether these have led to improvements in learning. The contribution made by extra-curricular activities (such as mentoring schemes and girls' clubs) to improving learning outcomes is clearer. These have raised girls' self-esteem and confidence in their academic ability, with some improvements in learning outcomes, particularly from activities that directly targeted girls' learning such as tutoring and accelerated learning programmes.

Five projects reported a backlash and resentment from boys who felt excluded from interventions. While this may partly be explained by boys being unwilling to give up gendered privileges, it also suggests that boys have many similar needs to girls that are not addressed by the GEC. As a response to these findings, several projects adapted their strategies to include boys in some of their activities.

Sustainability

Without additional funding from the next phase of GEC, it is highly unlikely that many SCW activities would have continued beyond the immediate short term. At the outset, projects were asked to report on the mechanisms they had put in place to enable marginalised girls to complete a full cycle of education. The FM reports that all projects have put these mechanisms in place. Projects' endline evaluation reports focused on the extent to which these mechanisms could continue to support girls without further support through the GEC. At this stage, projects are reliant on schools, teachers and community volunteers to continue to implement activities without further support. While there is evidence of commitment to do this, in the longer-term it is unlikely that this would have been possible, particularly for approaches and groups that are relatively new.

There has been some advocacy success by projects, and success in incorporating GEC materials into curricula and teacher training. These are important achievements given that many projects' advocacy activities have been constrained, or started late. Training (such as cascade training) undertaken by projects to facilitate the continuation of activities has also been carried out, but too late and there is little evidence of its use to develop the capacity of relatively new groups to continue delivering project activities. All projects have engaged with government at national, regional or local level, with evidence of some projects benefiting from favourable policy environments. This is evident in the SCW exceeding its match-funding target by 160 percent. However, government buy-in to actually take over SCW project activities has generally not been achieved, and government funding to sustain projects' activities does not appear to have been committed.

Value for Money

SCW projects prioritised their investments in interventions that addressed the most important barriers to girls' education, i.e. poverty and school-related barriers. Expenditure on economic interventions addressing poverty-related barriers and school-related interventions are associated with delivering the most learning beneficiaries. Amongst these, teacher training and support and school management and governance interventions are the most cost-effective at £64.70 and £56.20 on average per learning beneficiary respectively[15]. Economic interventions addressing poverty-related barriers and interventions designed to improve infrastructure and access to resources are the most expensive, at £123.90 and £104.60 per learning beneficiary respectively.

Among those interventions that delivered a substantial number of learning beneficiaries, extra-curricular activities and non-formal education seem to be the most cost-effective in delivering improvements in girls' learning. This finding is supported by projects' evidence that activities aimed at improving learning directly, such as tutoring clubs or accelerated learning programmes, have had a larger positive effect on girls' learning than 'indirect' activities such as scholarships or community-based activities focused on the benefits of staying in school.

The average unit cost per beneficiary for numeracy is about two to three times that of literacy across all types of intervention, which is consistent with the differences in the impact of the SCW on numeracy compared to literacy.

Despite benefiting a substantial number of marginalised girls, we found the SCW has not yet delivered the amount of improvement needed in literacy and numeracy to enable girls to progress through school and learn at the right

---

[15] Unit cost is calculated by dividing total expenditure per intervention by each project's learning impact (average of literacy and numeracy) relative to 0.5 Standard Deviations – each project's baseline to endline target – and then dividing by the total number of beneficiaries reached. Our value for money analyses excludes STC (Ethiopia), whose impact data is inconclusive, and RI (Somalia), who had a negative impact in both literacy and numeracy. These projects should in theory have 'infinite' unit costs, and were therefore omitted so as not to skew our calculations. Given this, average unit costs across SCW are understated, both at the aggregate and unit cost levels.

pace. More time may be needed for some interventions to deliver an increase in literacy and numeracy levels, but this finding also suggests that further changes in project designs and /or delivery processes are needed.

Key conclusions

Girls' learning levels were extremely low at baseline. SCW projects reached and benefited a substantial number of marginalised girls, but did not manage to impact on learning at a sufficient magnitude across all contexts. Girls in the communities targeted by the SCW are still falling substantially behind international benchmarks for reading fluency, which is likely to hinder their ability to learn effectively as they progress through school. Despite the absence of similar benchmarks for numeracy, the SCW results show that projects found it even harder to improve numeracy compared to literacy.

The EM data and analysis captures the effect of SCW projects on their target communities as a whole, and does not demonstrate any significant impact over and above the control group in any of the main GEC outcomes (i.e. attendance, literacy and numeracy). Projects' results suggest that the overall impact of the SCW has not been large and consistent enough across the window to be observed at the community level. Out of 14 SCW projects, eight met their literacy targets between baseline and endline, while four met their numeracy target. This does not mean that projects did not have a positive impact on their target populations, but many projects lacked the size and/ or magnitude of effect to enable them to achieve their targets. It is possible though that learning gains of projects' activities may not yet be fully realised within the relatively short timescale of the GEC.

SCW projects found it harder to improve girls' numeracy. Teaching numeracy requires specific "hard" skills that are not necessarily required to teach literacy. Our analysis shows that higher attendance in school is correlated with improvements in literacy to a much larger extent than in numeracy. Improving numeracy may require different types of interventions to those designed to improve literacy, potentially including higher investments in teacher training and improved teaching methodologies that are focused on numeracy.

Some subgroups face specific challenges and varying levels of marginalisation across contexts. Except for out-of-school girls, it was not a programme requirement to track these groups as specific cohorts, making it impossible to measure the impact of the SCW on their outcomes.

The effect of the SCW on specific subgroups of girls within the projects' target populations is unclear. Although there is evidence from the EM data that some activities successfully targeted girls who were most in need, this does not seem to have translated into better learning gains for these girls. The absence of an effect on these girls is largely because SCW projects usually considered their target population as a whole and did not differentiate between the varying needs of different subgroups within this population (e.g. orphaned girls, young mothers, girls with disability, etc.). When they did so, projects did not track the outcomes of these subgroups specifically. Even when subgroups can be identified in projects' data, associated sample sizes are usually not large enough to draw any meaningful conclusions about their learning progress. Besides, there is no consistent definition of subgroups across the SCW and projects often ended up using their own definitions, which makes comparisons at the window level even more difficult. Tracking large enough subgroup samples for evaluation and research purposes can be prohibitively expensive. However, for design, implementation and impact purposes, as much as for evaluation purposes, it is important that projects find cost-effective ways of identifying, diagnosing and addressing the different needs of the subgroups of marginalised girls that they are supporting.

Economic interventions have helped girls enrol and stay in school, but the assumption that this in itself will lead to improvements in girls' learning is highly dependent on the quality of teaching. There is a limit to the impact that SCW projects working at a local level, can have on households for whom poverty above all else is the binding constraint to girls' education.

Evidence from the EM data and projects' reports suggests that households generally have positive attitudes towards girls' education, but are unable to send girls to school because they do not have enough income to meet their basic needs. Scholarships, bursaries and stipends all helped households to offset the direct and indirect costs of education enabling them to send their girls to school. Sanitary wear and menstrual supplies provided by projects also encouraged more regular attendance. There is evidence that these interventions may be linked to

improvements in learning outcomes in countries where education systems are sufficiently developed and resourced to support the assumption that increased enrolment and attendance would naturally lead to improved learning.

Poverty though is a structural barrier to girls' education and underlies a range of different demand-side factors that prevent girls from accessing school and learning. The EM's evidence suggests that projects have had limited success in reducing the amount of time girls spend on household chores, or the effects of early marriage and pregnancies, which are often used by households as practical strategies for escaping poverty[16]. There seems to be a limit to the impact that SCW projects can have on households' ability to pay for schooling when poverty, above all else, is the binding constraint to girls being able to access a quality education.

A range of school-based interventions, such as infrastructure improvements, education materials, and literacy and numeracy clubs, have helped girls make substantial learning gains, but these gains have been limited by systemic constraints in the supply of education.

Problems with the supply of education infrastructure, learning materials and teachers are all symptomatic of education systems that lack the capacity and resources to provide a consistent quality of education. SCW projects are understandably tackling these issues at the school level where these symptomatic effects are most profound. However, there is little evidence that interventions are strategically informed by a holistic diagnosis of the education sector and the government and non-governmental systems, beyond the relatively short three-year life of the projects. While a diagnosis of education and school systems may have informed GEC project designs at the start, the evidence suggests that their achievements at endline are constrained by a severe lack of resources and capacity in these systems. As many SCW projects transition to the next phase of the GEC, they need to use their data and evidence to reassess the effectiveness of each intervention in the context of the institutions in which they are embedded. If they do this, projects will be in a better strategic position to pre-empt and avoid situations where, for example, new sanitation facilities lack water, or where projects are confronted by severe systemic problems.[17]

The lack of evidence in project reports of the effectiveness of different types of interventions hindered our capacity to evaluate what worked well (or less well) in improving girls' learning outcomes.

Projects need to more systematically capture and present data and analysis in their reports that shows how, why and to what extent their activities contributed to reducing barriers to girls' education, and how changes in these barriers specifically link and contribute to learning outcomes. It is important for projects to understand and evidence, in a more granular way, what specifically worked well (or not) to inform course corrections and identify which activities should be sustained. As recommended above, a greater focus on identifying and tracking the characteristics of different subgroups will help projects do this. Similarly, identifying and evidencing the intermediate steps between outputs and outcomes as an integral part of their M&E frameworks would help explain how and to what extent some interventions are having a greater effect on learning than others. The Terms of Reference for projects' external evaluators should be more prescriptive, including requirements to: use a standard household survey template; link household survey data with learning outcome data; and collect specific disaggregated data about key subgroups (which was not a GEC requirement for projects).

Without the support of the successor to the GEC programme, it is highly unlikely that many SCW activities would have continued, given the severe lack of resources available at all levels of the education systems in which projects operate.

Without the support of the successor to the GEC programme, it is highly unlikely that many SCW activities would have continued as they stand now and very few would have continued beyond the immediate short term. This is not surprising, as the business case recognises that for all windows (not just the SCW) there was a risk that there would be a trade-off between sustainability and achieving short-term results. Many projects' sustainability strategies were overly reliant on communities, schools or government ministries to take over responsibility for continuing activities without much further support or funding. There is little evidence that this can happen given the severe lack of resources available at all levels of the education system.

---

[16] Terry et al (2016), 'Narrow Windows, Revolving Doors – GEC Thematic Research Report', Coffey International Development
[17] Reported in the endline evaluation reports of BRAC (Afghanistan), AKF (Afghanistan) and ChildHope (Ethiopia).

Although sustainability strategies were developed too late in projects' lifecycles, the underlying barrier is nevertheless a pervasive lack of resources within the target communities, schools and education systems. Sustainability strategies should explicitly recognise the reality of trade-offs between addressing symptomatic effects to achieve short-term learning gains versus what might be considered a longer-term objective of addressing the structural causes of poor education outcomes.

The SCW has improved the literacy and numeracy outcomes of a large number of marginalised girls living in challenging conditions while simultaneously facing under-resourced school and education systems. In this context, projects need to significantly adapt their designs and delivery processes to increase their effects on girls' literacy and numeracy, to enable them to effectively progress through their education.

Despite working in and with under-resourced school and education systems, the SCW has improved the literacy and numeracy outcomes of a substantial number of marginalised girls living in a range of challenging conditions. However, girls' baseline learning levels in the communities targeted by the SCW were seriously low and they still are at endline. SCW projects need to deliver a greater magnitude of improvement in girls' literacy and numeracy levels, to improve the effectiveness and overall value for money of the outcomes currently being achieved. It may be that more time is needed for some interventions to deliver their full potential, but the overall findings from this evaluation suggest that significant changes in some project designs and delivery processes are needed to achieve the literacy and numeracy gains that these marginalised girls need to progress with their education.

1 Introduction

1.1 Background and context of the GEC

1.1.1 Context and rationale

Changes in global problems affecting the education of marginalised girls

Every child has the right to a basic education. Educating girls, especially to secondary level, delivers significant economic, health and social benefits to the girls themselves, their families and the wider community. Girls who complete secondary school tend to have fewer children over the course of their lives, higher wages, and a higher life expectancy compared to girls who have only completed primary school.[18] A recent report by the Education Commission suggests that one US dollar invested in an additional year of schooling in low-income countries, particularly for girls, creates earnings and health benefits of US$10, and around one-third of the decrease in adult mortality since 1970 comes from improvements in the education of girls and young women.[19]

From the outset of the Girls' Education Challenge (GEC) Fund in 2012, DFID was working towards the Millennium Development Goals (MDGs) and the international targets agreed by the United Nations (UN) to halve world poverty by 2015. Progress on girls' education was critical to the achievement of MDGs 2[20] and 3[21], which specifically relate to universal primary education and gender equality. By the time that the final Education for All (EFA) Global Monitoring Report (2000-2015) was published, more girls were in school, stayed there longer, and learned more whilst they were there[22].

The last decade has seen the world approaching universal primary education coverage with a majority of children, both boys and girls, entering primary education in most countries around the world. Although primary school enrolments for girls have improved along with boys, school completion rates remain equally low for both boys and girls. Around 264 million children and young people were estimated as being out of school by the end of 2015 – this includes 61 million (9%) of children of primary school age (6-11 years), 62 million (16%) children of lower secondary school age (12-14 years) and 141 million (37%) children of upper secondary school age (15-17 years)[23]. In secondary school, the differences between boys' and girls' participation rates are significant. While global averages may suggest that gender inequalities are small, at regional and national levels, large disparities exist within countries, with poor rural girls experiencing worse educational outcomes, even at the primary school level. Girls are more likely to remain completely excluded from the formal education system than boys. Across Sub-Saharan Africa, 9 million girls will never have the opportunity to learn to read or write in school (compared to 6 million boys), and in South Asia, four out of five out-of-school girls will never attend school (compared to two out of five out-of-school boys).

Improved access to education has not resulted in improvements in learning for many children. Levels of learning remain appallingly low, even for those enrolled in school[24] - approximately 250 million children cannot read, write or do basic maths, including more than 130 million children who do not have basic literacy skills, despite being enrolled in primary school. While significant challenges persist in terms of access to education, global development organisations recognise that improvements in the quality of learning in schools are urgently needed. The

---

[18] United Nations Girls' Education Initiative (2014): *Accelerating Transition of Girls to Secondary Education: A Call for Action.* Discussion Paper, United Nations Children's Fund (UNICEF), New York, published online at: http://www.ungei.org/resources/files/2014-04-GPE-UNGEI-Accelerating-Secondary-Education-Girls.pdf.

[19] The International Commission on Financing Global Education Opportunity (2016): *The Learning Generation. Investing in education for a changing world.* Available at: http://report.educationcommission.org/wp-content/uploads/2016/09/Learning_Generation_Full_Report.pdf

[20] Goal 2: *Achieve universal primary education*; Target 2.A: Ensure that, by 2015, children everywhere, boys and girls alike, will be able to complete a full course of primary schooling.

[21] Goal 3: *Promote gender equality and empower women*; Target 3.A: Eliminate gender disparity in primary and secondary education, preferably by 2005, and in all levels of education no later than 2015.

[22] EFA (2015), 'Global Monitoring Report: Education for All 2000-2015, Achievements and Challenges', UNESCO

[23] UIS and GEM (2017), 'Reducing global poverty through universal primary and secondary education. Policy paper 32 /Fact Sheet 44. Paris: Global Education Monitoring Report. Available at: http://unesdoc.unesco.org/images/0025/002503/250392E.pdf

[24] UNESCO (2014), 'Teaching and Learning: achieving quality for all'. EFA Global Monitoring Report, Paris: UNESCO

Sustainable Development Goals (SDGs) explicitly recognise the need to address not only persisting gaps in access, but also gaps in the equity, inclusivity and quality of education worldwide.[25]

It is estimated[26] that over 36 per cent of children who are out of school (28 million of primary school age) live in areas of conflict – this includes countries targeted by the GEC's Step Change Window projects such as Afghanistan, Somalia, Kenya and DRC. In fragile and conflict-affected states and in crisis situations the provision of basic education services becomes difficult or impossible.

Persistent under-investment in education

A key rationale[27] in 2011-2012 for DFID's investment in the GEC was that traditional Overseas Development Aid (ODA) to education had stagnated and, given the global financial situation and shifting development priorities, may even decline. Now, in 2016, under-investment in education persists. In 2013 aid data released by the OECD's Development Assistance Committee (DAC) showed a decline in education aid for the third consecutive year, with basic education (primary and secondary) suffering the greatest decline. While total ODA rose by 11 per cent in 2013, aid to basic education declined by 7 per cent[28]. The UN Educational Scientific and Cultural Organisation (UNESCO) estimates that more than double the current levels of spending would be required to achieve the SDG education targets by 2030. Education in humanitarian and conflict-affected settings continues to receive a relatively small proportion of the humanitarian budget (less than 2 per cent), which prevents those children who are most marginalised from accessing a quality education.

Changes in the global policy response to education

The GEC Step Change Window Baseline Report[29] was published in 2015, which is also the year that marked the end of the MDGs, and the adoption of the SDGs. Goal 4 of the SDGs seeks to 'ensure inclusive and quality education for all and promote lifelong learning'. This goal recognises that major progress has been made towards increasing access to education at all levels and increasing enrolment rates in schools, particularly for women and girls, but a greater focus is needed on the quality of education provided to enable effective learning outcomes to be achieved. Goal 5 aims to 'achieve gender equality and empower all women and girls'. This entails tackling discrimination against women and girls, including issues of: violence and sexual violence; early and forced child marriage; female genital mutilation/cutting (FGM/C); effective participation and equal opportunities for leadership at all levels of decision-making; and access to health services, in particular reproductive health services.

As well as changes in the global policy response to education, there have also been changes in global funding for education – for example, the Education Cannot Wait Fund. The Fund was launched in 2015 and is hosted by UNICEF. The Fund responds directly to the SDG commitment of achieving a quality education for all and leaving nobody behind. This was the second global fund to prioritise education in humanitarian settings, following on from UNICEF's Education in Emergencies and Post-Crisis Transition Fund 2007 – 2009.[30] In 2016, the Education Commission also proposed an International Finance Facility for Education (IFFEd)[31] that could mobilise an additional US$10 billion annually for education by 2020. The model brings together public and private donors, alongside international financial institutions like the World Bank and the regional development banks to create low-interest finance packages for lower-middle-income countries linked to countries increasing their own level of investment in education and carrying out education sector reforms. An overview of other girls' education programmes is given in Annex C.

---

[25] United Nations Economic and Social Council (2016): Progress towards the sustainable development goals. Report of the secretary-general. E/2016/75, available at http://www.un.org/ga/search/view_doc.asp?symbol=E/2016/75&Lang=E.

[26] EFA (2015), 'Global Monitoring Report: Education for All 2000-2015, Achievements and Challenges', UNESCO

[27] DFID (2012), Girls' Education Challenge, Business Case Version 4; London: DFID

[28] Steer and Smith (2015): "It's time to Reverse Declining ODA to Education", Brookings, published online at: https://www.brookings.edu/2015/01/12/its-time-to-reverse-declining-oda-to-education

[29] Coffey (2015) 'Baseline Report – Step Change Window' https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/425360/Step-change-window-full2.pdf

[30] https://www.unicef.org/emergencies/files/FINAL_PUBLIC_Consolidated_Netherlands_Donor_Report_2009.pdf

[31] The Education Commission (2016) The Learning Generation. Investing in education for a changing world. Education Commission

The Girls' Education Challenge Fund and its extension

In 2012, the Department for International Development (DFID) launched the £355 million Girls' Education Challenge Fund (GEC). The GEC intends to support up to a million of the world's most marginalised girls to improve their lives through education. The GEC has provided this support through three separate funding windows:

- the Step Change Window (SCW);

- the Innovation Window (IW); and

- the Strategic Partnerships Window (SPW).

All GEC projects work towards the same high-level GEC outcomes of improved enrolment, retention (measured through attendance), literacy and numeracy for marginalised girls. However, each window has distinctive features and a specific focus.

The current GEC programme ended in March 2017, with individual projects' contracts ending at various points between early 2017 and March 2017. At the Girls' Education Forum in London in July 2016, DFID announced it would provide a further £100 million of funding to the GEC[32]. This funding will continue its support for: the one million marginalised girls supported by the GEC, enabling them to progress, transition through school phases and complete a cycle of education; and to help an additional 175,000 of the poorest and most marginalised girls receive a quality education – the Leave No Girl Behind initiative[33]. This is a new initiative that will support:

- interventions providing literacy, numeracy and skills relevant for life and work to adolescent girls aged between 10 and 19 who have never attended or have already dropped out of school; and

- girls located in one of the countries where DFID works and who are highly marginalised – girls who experience complex marginalisation because of their circumstances, including orphans, married or young mothers, girls with a disability, nomadic girls, refugees, those from the poorest communities and those with no access to education.

1.1.2 GEC theory of change and assumptions

The EM produced a high level Theory of Change (ToC) for the GEC as part of the GEC Evaluation Strategy[34] produced in 2013. This theory of change and DFID's Business Case for the GEC assumes that there are educational barriers that affect boys and girls, but that girls face a number of additional, gender-specific challenges which put them at a disadvantage in comparison with boys. The ToC for the GEC Evaluation Strategy showing the links between different types of barriers and the GEC's outcomes is presented in Annex A.2.

The overarching ToC sets out the problems and barriers that hinder girls from enrolling, attending and learning in school. Problems identified include: economic barriers such as school fees, opportunity costs, and the cost of school materials; social and cultural barriers such as restrictive views about girls' education and the role of women and girls; educational barriers such as a lack of female teachers and poor teaching; logistical barriers such as a lack of appropriate school facilities and distance to school; and institutional or political barriers such as lack of equity in the public service provision, or a lack of influence of and accountability to marginalised groups.

Through the challenge fund design, the GEC encouraged organisations to develop their own theories of change and intervention mechanisms to address (some of) these barriers in specific contexts and for specific populations of marginalised girls. The ToC assumes that by tackling these barriers, projects will help to improve girls' enrolment, retention (measured through attendance), literacy and numeracy in school and contribute to an overall impact of improved life chances for marginalised girls.

---

[32] For more details see DFID's press release on the extension of the GEC, at: https://www.gov.uk/government/news/britain-to-help-175000-girls-in-worlds-poorest-countries-get-an-education.
[33] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/559123/leave-no-girl-behind.pdf.
[34] Coffey (March 2013). The ToC can be found in Annex A in the 'GEC Evaluation Strategy'. The GEC business case can be found at: https://devtracker.dfid.gov.uk/projects/GB-1-202372/documents.

1.1.3 Summary of GEC projects and interventions

Fifteen projects were awarded funding of up to £30 million per project through the GEC SCW to quickly and effectively expand education opportunities for girls at primary and secondary school levels. They operate in nine countries: Afghanistan, DRC, Ethiopia, Kenya, Mozambique, Sierra Leone, Somalia, Tanzania and Zimbabwe.

All fourteen[35] projects that were operating at the time of the endline evaluation developed holistic theories of change that aimed to address multiple barriers to girls' education at the level of the individual girl, the household, the community, and the school. Projects tackle these barriers through a wide range of interventions. The GEC Fund Manager (FM) has produced an intervention mapping that is summarised in Table 1.

ACTED (Afghanistan) stopped their activities at the time of the midline evaluation because of concerns around insecurity in their target areas, partly related to Taliban advance. The evaluation report they submitted at midline is therefore considered their endline report and used as such in our analysis.

Plan (Sierra-Leone) activities were considerably disrupted by the Ebola outbreak in 2014 with all schools in the country closing for nine months. The project did not carry out any evaluation activity at midline but tried to follow up their cohort at endline[36], and assessed learning progress of girls in their target communities. See the Plan 1 endline evaluation report for further details.

The FM has distinguished "core" interventions from more peripheral interventions that are considered less central to the project's intervention strategy, based on the FM's knowledge of each project's activities. **According to this mapping, all projects carry out economic interventions** to offset the cost of education (e.g. bursaries, cash transfers, in-kind support or loans); **run activities to improve school infrastructure and resources** (e.g. improving classrooms, providing textbooks and materials, improving sanitation facilities, etc.); **provide teacher training and support** (e.g. in literacy and numeracy or gender-sensitive pedagogy); **work with communities** (e.g. through media campaigns, cooperation with parent and women's groups, engagement of faith leaders, etc.); and **strengthen school governance and management structures.**

In addition, **some projects** carry out additional **activities to empower girls** and raise their self-esteem, to **tackle marginalisation** (e.g. by addressing barriers related to disability, or issues of cultural or linguistic exclusion), and to **reduce violence** in school or in the community.

A full overview of the activities that each project is delivering, as well as a more detailed intervention mapping, is provided in Annex B. A discussion of each project's target group(s) and the extent to which these have been reached by the time of the endline evaluation is provided in Section 3.1.

Table 1: Overview of SCW projects' interventions

Lead Partner	BRAC	AKF	Acted	IRC	STC	ChHp	WUSC	CfBT	STC	Plan	RI	CARE	WV	Camf
Country	Afg	Afg	Afg	DRC	Eth	Eth	Ken	Ken	Moz	Sie	Som	Som	Zim	Z-T[(1)]
ECONOMIC — Economic interventions offsetting the cost of education: Bursaries, Cash Transfers, Income-generating activities, In-kind support (school kits, menstrual supplies), Loans and savings.														
	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
INFRASTRUCTURE & RESOURCES — Infrastructure and resources for schooling: School and classroom building/ improvement; Technology in classroom; Textbooks & Learning materials; Toilettes & WASH facilities.														
	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Teacher training and support: Formal pre-service teacher training; Gender responsive pedagogy; Inclusive classroom strategies; Literacy and numeracy; Peer support and mentoring; Skills training.														

---

[35] Please note that only fourteen projects were active at the time of the endline evaluation. The GEC grant held by BRAC for its project in Sierra Leone ended in 2015. BRAC was put on a Project Improvement Plan (PIP) shortly before the State of Emergency due to the Ebola outbreak was called, due to concerns with the design and implementation of the project as well as the capacity of the management to address these concerns. The challenges of the Ebola epidemic exacerbated the issues faced by the project which was eventually closed by recommendation from the Fund Manager and approval from the DFID GEC Steering Committee. The beneficiaries from the BRAC project were taken on by Plan in a newly created project "Plan 2", which was followed through to endline. More details are given in the Plan 2 endline evaluation report.
[36] Only 21% of baseline girls were successfully recontacted at endline (See Table 5 on Attrition rates by SCW project).

TEACHER TRAINING															
	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	

COMMUNITY BASED
Community-based interventions: Adult literacy; Community gatherings; Household-level visits and support; Media (radio, TV, advertising); Parents' and women's groups; Working with faith groups and traditional leaders; Working with men and boys.

	✓	✓	✓	✓	✓	✓	✓	✓	✓	✦	✓	✓	✓	✓

EXTRA-CURRICULAR & NON-FORMAL EDUCATION
Extra-curricular activity & non-formal education: Life skills (incl. SRH); mentoring (e.g. peer support); mixed sex or boys' clubs; non-formal / alternative education; tutoring (e.g. homework clubs); vocational training & economic empowerment.

	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✦	✓	✓

SCHOOL MANAGEMENT & GOVERNANCE
School management and governance interventions: Community and private schooling provision; Technology for school management; Working with local or national education authorities; Working with SMCs, PTAs and other stakeholders.

	✓	✓	✓	✓	✓	✓	✓	✓	✓	✦	✓	✓	✓	✓

EMPOWERMENT & SELF-ESTEEM
Empowerment and self-esteem: Activities that promote girls' voice and participation; Mentoring; Role models (older girls, female teachers, parents); Safe spaces.

	✓	✓	✦		✓		✓		✦	✦	✓	✓

MARGINALISATION-RELATED
Marginalisation-related interventions: Interventions in remote or nomadic locations; Interventions addressing cultural/ linguistic exclusion; Interventions addressing disability; Interventions with other marginalised groups.

	✦			✓	✓	✓	✦		✓		✦	✓	✓

VIOLENCE-RELATED
Violence-related interventions: Addressing abuse from adults in charge; Addressing child marriage and FGM; Addressing corporal punishment; Addressing peer violence; Child protection policies development in schools.

	✦	✓	✦		✓	✓		✓		✓	✓	✓	✦

Key: ✓ indicates that an intervention of this type is at the core of the project's intervention strategy.
✦ indicates that an intervention of this type is used by the project, but is not a core activity.
(1) Z-T: Zimbabwe and Tanzania.
Source: Adapted from the intervention mapping produced by the FM.

As part of the challenge fund design, each project designs and delivers its own set of activities and maintains relationships with local, regional and national education authorities and stakeholders. Most projects[37] are delivered through a consortium of national and international partners, led by an international NGO that is accountable for the grant agreement held with the FM on behalf of DFID. Cooperation between projects working in the same country or region is encouraged but is not an explicit requirement of the GEC.

> **!** For simplicity, readability, and consistency with the SCW baseline and midline reports, we refer to SCW projects by the name of the Lead Organisation throughout this report. However, it is important to keep in mind that most projects are being delivered by a consortium of partners who should be recognised equally for their contribution. More information on national and international project partners can be found in the projects' individual endline evaluation reports.

As shown in Table 1, all SCW projects take a 'holistic' approach to addressing barriers to girls' education by combining interventions at the household, community, school, and governance levels. This approach reflects the projects' understanding at the start of the programme that to tackle barriers to girls' access to a quality education a holistic approach was needed. Differences in approaches or intervention designs do not show at this high level as they only emerge when looking at the specific activities that projects are delivering (see Annex B).

As projects use holistic intervention designs and carry out a bundle of different activities in the same communities, it makes it very difficult to distinguish the effectiveness of different types of activities using quantitative methods, as we cannot isolate the effect of one intervention from the contributions of another. When planning for monitoring and evaluation (M&E), projects were not required to design impact evaluations that could isolate individual intervention

---

[37] 11 out of 14 SCW projects are working with a national or international delivery partner. The other projects do not have a key partner organisation.

effects, which would have been practically difficult and very costly. This would have been more feasible, had projects chosen to focus on single interventions, but this was not a requirement of the GEC.

Figure 1 shows the budget spent by SCW projects on each intervention type since baseline, except Plan (Sierra Leone) for which no data was available. The total budget across the SCW amounted to £220m, with Economic interventions representing the highest share of projects' budgets, before Infrastructure and resources and Teacher training and support.

Figure 1: Budget spent by all SCW projects by intervention type (million GBP)
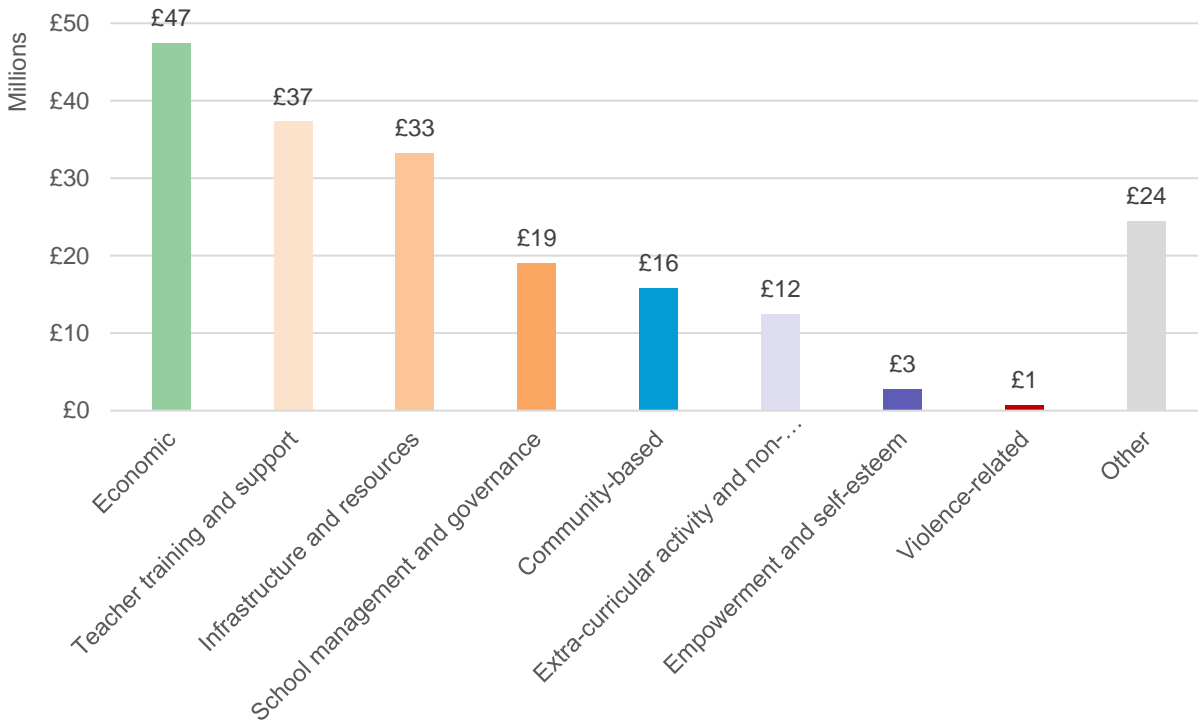
The light purple bar denotes "Extra-curricular activity and non-formal education". The "Other" category represents operation and overhead expenses, human resources, salaries and administrative costs, M&E and other types of activities. It is worth noting that some of the budget categories shown in Figure 1 may be overestimated or underestimated compared to the types of intervention actually run by the project, because this classification involved some subjectivity, especially in cases of overlaps between two categories. When a type of intervention could fall into two different categories, we favoured its type over its target group or its theme. For instance, "Teachers trained on Safe School Approach" was classified under "Teacher Training" rather than "Violence-related" intervention. The "Violence-related" budget category is therefore underestimated compared to other categories.

1.2   Governance, purpose and scope of this evaluation

1.2.1   Governance of this evaluation

In 2012 DFID appointed Coffey, in partnership with RTI International and ORB, as the Evaluation Manager (EM) of the GEC. We are responsible for designing and implementing the GEC monitoring and evaluation (M&E) framework to assess the effectiveness and impact of the programme as a whole. We also generate and share lessons learned to inform the ongoing design and development of the GEC programme and wider DFID programming. Annex D provides an overview of the roles and responsibilities of the different EM consortium partners.

We closely collaborate with the GEC Fund Manager (FM) (a consortium led by PwC) to support projects in collecting data in line with the evaluation requirements, and in reporting results with a maximum level of consistency across the fund. The FM has played a key role in developing M&E processes and requirements at the

project level, and in managing relationships with projects. Annex D shows the M&E activities carried out by the FM in the GEC.

The 14 projects funded through the SCW that are the focus of this endline evaluation have been responsible for developing their own project-level M&E frameworks. They have each been required to contract an external evaluator who collects data and assesses their progress and performance independently at the project level. The FM and EM reviewed and quality assured the research instruments and reports produced by projects and their external evaluators. Our intention, as the EM for the GEC, was to aggregate project datasets for meta-analysis. Unfortunately, due to the insufficient level of consistency across project data (inconsistent surveys, sampling designs, learning tests, data quality, etc.) this has not been possible (see Section 2.5.2 about the limitations of SCW projects' data).

1.2.2   Purpose of the GEC endline evaluation

The overarching purpose of the GEC Evaluation Strategy is to produce reliable evidence of the programme's effectiveness and impacts that DFID, the FM and projects can use to inform improvements during the programme's lifetime, as well as future programme design. In particular, it is expected that DFID, the FM and projects will use the findings and lessons learned from this evaluation to inform the successor programme to the GEC and its new Leave No Girl Behind window that DFID announced in July 2016.

DFID always envisaged that the programme evaluation should generate transferable lessons about what works, what does not, where and why in delivering girls' education outcomes for a wider audience including its partners, governments of GEC countries, and other policy-makers. In addition, the EM has developed, with DFID, a specific communication plan for the findings of this endline evaluation that will target key education partners of DFID, including UNICEF, GPE, UNGEI, USAID, GPE, UNESCO and the World Bank. Activities in the communication plan will be delivered in late 2017 and early 2018 following the publication of this report.

This endline evaluation also serves an important accountability purpose by providing reliable information about the effectiveness and impact of the SCW projects at the end of the three-year implementation period. It follows the baseline evaluation (2013/2014), and the midline evaluation (2015/2016).

1.2.3   Communication plan

The first phase of the GEC programme came to an end in March 2017. The GEC Knowledge Management Working Group was led by the FM and performed a coordinating role in identifying and facilitating opportunities to communicate and disseminate learning across the GEC programme and beyond, to inform wider policy-making and programming. As the EM, we were an active member of this working group together with DFID. With the FM now focusing on supporting the programme to the next phase of the GEC (i.e. GEC Transitions and accompanying Leave No Girl Behind window), this working group has now ceased operating. However, the FM is still committed to actively supporting the communication and dissemination of the results from this evaluation. This also does not affect DFID's commitment to sharing these results within DFID and with its partners. The FM, DFID and EM continue to meet on a weekly basis. These meetings will be used to develop and implement the communication plan for these evaluation results as effectively as possible.

The final SCW Endline Evaluation Report will be completed for publication by the end of November 2017. The EM's contract ends in January 2018, which allows a limited amount of time for us to actively communicate the evaluation results as part of this evaluation programme, although we are committed to supporting further communication and dissemination by DFID as far as possible after the end of the programme.

The findings from this endline evaluation will be communicated to DFID, the FM, projects and other stakeholders through a combination of: face-to-face presentations; webinars; seminars; and through the publication and dissemination of the SCW Endline Evaluation Report and an Evaluation Briefing Note that summarises the key findings from the report in two to four pages for publication by DFID on the GEC website[38]. A similar set of communication activities were delivered to disseminate the results from the SCW Midline Evaluation Report, which proved effective. Prior to completion of this report, together with the FM, DFID and three GEC projects, we

---

[38] https://www.gov.uk/guidance/girls-education-challenge

presented the emerging findings from this endline evaluation and those emerging from our evaluations of the IW and SPW at the biannual UKFIET conference[39] in September 2017.

1.2.4   Scope of the GEC SCW endline evaluation

The endline evaluation aims to answer the following research questions:

- To what extent have target girls and their communities been reached by SCW project interventions?

- To what extent have SCW projects improved girls' enrolment, attendance, and learning?

- To what extent are SCW project interventions addressing key barriers to girls' education and with what effects?

- What type of interventions work, in what context, and for whom?

The SCW endline evaluation focuses on changes in outcomes (i.e. attendance and learning) and steps towards the achievement of outcomes (i.e. barriers to girls' education). Reporting on outputs and progress in delivering GEC activities is covered through the FM's performance reporting, and is not within the scope of this evaluation. A discussion of progress against output targets can be found in each project's endline evaluation report.

The SCW endline evaluation focuses on answering questions about the effectiveness and impacts of the SCW projects, but also explores the value for money (VfM) delivered across the funding window, as well as the potential sustainability of programme activities. The EM's assessments of VfM and sustainability rely on data and information provided by the FM and in projects' endline evaluation reports.

---

[39] UKFIET, The Education and Development Forum – https://www.ukfiet.org/conference/

2 Evaluation approach and methodology

2.1 Overview of the GEC evaluation strategy

For the SCW, the GEC Evaluation Strategy sets out a combination of EM-led and project-led evaluation activities that generate evidence at the project and window levels. The EM and a majority of projects (see Table 2) collected data in intervention areas and in matched control areas to enable a counterfactual evaluation of impact. This means that impact of the SCW is being measured as the net change observed in the treatment group over and above the change measured in the control group, using a difference-in-difference indicator.

2.1.1 Overview of the SCW projects' evaluation design

SCW projects assess the impact of their interventions on their specific target groups (see Section 3.1 for an overview of each project's target groups). Projects generate findings about what works, what does not, and why at the project level, draw lessons learned about their theories of change, and reflect on possible improvements to their project design. The project-level evaluations include the following activities:

- Commissioning an external evaluator to collect data at baseline, midline and endline and producing evaluation reports that comply with a template provided by the FM and EM.

- Collecting a combination of quantitative and qualitative data in intervention and control communities (or schools), including a longitudinal household survey. SCW projects are required to use the survey questionnaire template provided by the EM to collect data consistently across the window.

- Testing literacy and numeracy using standardised international tests, and conducting appropriate statistical analysis to report on changes in learning outcomes.

Six SCW projects work in fragile and conflict-affected contexts in Somalia, Afghanistan, and Kenya[40]. During preparations for baseline data collection it became clear that the use of control groups in these unstable environments could increase the risk of conflict or escalate existing tensions between conflicting clans, ethnic groups, and power elites by allocating interventions to some communities but not others. The risk of violence against treatment schools, children and their families, as well as project teams and enumerators, was considered high. In addition, local governments disapproved of the use of control groups, which further endangered the smooth implementation of programme activities.

In light of these significant challenges, DFID and the FM considered that the use of control groups would go against the principle of "Do No Harm" and decided to drop control groups in Somalia, Afghanistan, and WUSC (Kenya) project areas. As shown in Table 2, the eight other SCW projects used a randomised control trial (three projects) or a quasi-experimental evaluation design including control groups (five projects).

Table 2: Evaluation designs and learning tests used by SCW projects to measure impact at endline

Endline	BRAC	AKF	ACTD	IRC	STC	ChHp	WUSC	CfBT	STC	Plan	RI	CARE	WV	Camf
	Afg	Afg	Afg	DRC	Eth	Eth	Ken	Ken	Moz	Sie	Som	Som	Zim	Z-T
Use of control groups	✦	✦	✦	RCT	QE	QE	✦	QE	RCT	QE	✦	✦	RCT	QE
Learning test used	EGRA/ EGMA	EGRA/ EGMA	ASER	EGRA/ EGMA	EGRA/ EGMA	EGRA/ EGMA	EGRA/ EGMA	UWEZO	EGRA/ EGMA	ASER	UWEZO	UWEZO	EGRA/ EGMA	National
Household-based or school-based survey?	Mix	Mix	School	School	HH	Mix	Mix	HH	HH	Mix	Mix	Mix	Mix	School
Out-of-school girl beneficiaries	No	No	No	Yes[41]	No	Yes	No	Yes	Yes	No	Yes	Yes	No	No
RCT	The project conducts a randomised control trial.													
QE	The project uses a quasi-experimental evaluation design.													
✦	The project uses a before and after design without control groups.													

---

[40] WUSC (Kenya) targeted girls in the Kakuma and Dadaab refugee camps and the local host communities.
[41] The sampling is predominantly school-based with a large proportion of beneficiaries (96%) being in school. However, a small cohort of 100 out-of-school girls were sampled outside school.

2.1.2 Overview of the EM evaluation design

The GEC EM assesses the overall impact of the SCW on girls who were aged 5-15[42] at baseline, and live in the catchment areas of SCW projects' treatment and control schools. The EM conducts longitudinal surveys and learning tests with selected girls at three points during the programme's lifecycle: baseline, midline and endline. Sampling girls in the community makes it possible to survey both in-school and out-of-school girls and their families, to measure enrolment rates, and to explore barriers and enabling factors influencing the girls' education outcomes that are situated within the family and community environment. For more details on the EM sampling approach, see Section 2.3.

The EM-led research complements the project-level evaluations by collecting harmonised data across the SCW to assess the programme's wider effects on target communities. At endline, this research included the following activities (see Annex E for more details):

- A longitudinal Household Survey (HHS) and a learning assessment with one selected girl in each surveyed household. The survey and learning tests are conducted at baseline, midline and endline. At endline, the sample includes 6,279 households and covers both intervention and control areas across the 14 SCW project areas[43].

- A follow-up School Visit Survey (SVS) to confirm the enrolment reported by the girls' caregivers in the household survey and gather some contextual information about the girls' schools, classrooms and teachers. The SVS has been conducted in all countries at endline except Afghanistan, Sierra-Leone and Camfed's project areas in Zimbabwe[44]. The SVS covered 3,107 girls at endline.

- Review and meta-synthesis of the endline reports, datasets and outcome spreadsheets submitted by SCW projects and collated by the FM.

The baseline and midline EM research activities also included qualitative In-Depth Interviews (IDIs) among selected caregivers, teachers and community leaders, as well as a School-Based Assessment (SBA) conducted across SCW project areas in DRC, Ethiopia, Kenya and Sierra-Leone. IDIs and SBA are not part of the endline EM evaluation. A discussion about the differences in EM research activities across different project areas at endline can be found in Section 2.5.1.

2.1.3 Changes to the EM evaluation design and methodology since baseline

The EM overarching evaluation design has not changed since baseline, and the fieldwork methodology has only seen minor changes. The endline research was conducted with the same longitudinal sample of girls and their households that was drawn at baseline and re-contacted at midline. EM endline attrition rates by SCW project area are shown in Table 4. New girls and households were only added to the sample to substitute respondents that could not be re-contacted at either midline or endline, or who refused to be interviewed again. In a few cases, an entire sampling point had to be substituted because the village or school catchment area originally surveyed had become inaccessible. Details are provided in Annex E.

The Ebola outbreak in Sierra Leone in 2014 caused Plan International to revise their project intervention strategy. An additional "Education in Ebola" element was funded and added to the Plan (Sierra Leone) project to support the government's response to the outbreak and respond to the ends of the target girls during the crisis. A special module on Education in Ebola was inserted into the EM midline and endline household surveys to assess the impact of the Ebola crisis on girls' education.

---

[42] Or 5-17 in Camfed project areas. See Section 1.3 in Annex E for more details on the EM endline research methodology. All other project areas were assigned a control group, including those located in Somalia, Afghanistan and Kenya refugee camps and host communities (WUSC), where control groups had been dropped from projects' research in application of the "Do No Harm" principle. Control groups were considered acceptable in the EM research as they were not perceived to be directly linked to any interventions.

[43] Please note this includes projects who have dropped control groups for their own evaluation activities since the EM and projects' household survey samples are independent from each other (see Section 2.3 for more details).

[44] No school-visit survey was carried out in Afghanistan at baseline to avoid any tensions that may have arisen in the run up to the national elections. The school-visit survey was not reintegrated in the midline and endline EM research designs. In Sierra-Leone, school research activities were cancelled in 2014 and 2015 due to the Ebola outbreak and were not reintegrated at endline. For Camfed, this is due to the alternate sampling procedure that was used in their target community areas (see Section 1.3 in Annex E).

2.1.4 M&E support provided to projects at endline

The FM supports projects with their evaluation and monitoring systems on a continuous basis, advises projects on changes to their evaluation design, and the adequacy of learning tests and analytical models. The FM also leads the quality assurance of projects' research instruments, datasets, and evaluation reports. The EM supported projects as follows at endline:

- Provided projects with a household survey template for endline, as well as a guidance package explaining how the survey has changed between midline and endline, discussing good practice in supervising fieldwork processes, advising on the process for dealing with survey attrition, and for merging and submitting household survey datasets to the FM and EM;

- Shared this guidance with projects and their external evaluators through webinars hosted by the FM;

- Developed a detailed project endline evaluation report template; and

- Reviewed and quality-assured projects' endline research instruments and evaluation reports jointly with the FM.

2.2 Sources of evidence

As described in Section 2.1 our SCW endline evaluation report draws evidence from a range of different data sources. Table 3 presents a simplified version of the GEC evaluation framework, showing how data sources are being triangulated to answer each evaluation question.

Table 3: Overview of the streams of evidence used to inform the analysis presented in this report

| GEC Evaluation Questions | Data sources used for analysis | | | | | |
| | Project | | | EM | | |
	Project Reports	Project Data	Outcome spread-sheets	Household Survey	Learning tests (EGRA/ EGMA)	School Visit Survey
Relevance: To what extent has the GEC reached marginalised girls? (Section 3.1)	✓		✓	✓		✓
Impact: What impact has the GEC had on enabling marginalised girls to be in school? (Section 3.2)				✓		✓
Impact: What impact has the GEC had on marginalised girls' learning? (Section 3.3)		✓	✓	✓	✓	✓
Effectiveness: What has worked, why and with what effects? (Section 3.4)	✓	✓		✓	✓	✓
Efficiency: To what extent does the GEC represent good value for money? (Section 3.6)	✓	✓				
Sustainability: How scalable and sustainable are the activities funded by the GEC? (Section 3.7)	✓					

We use quantitative data from the EM surveys as well as the project datasets and outcome spreadsheets to assess the magnitude of change in education outcomes and educational barriers that projects across the SCW have achieved since baseline and midline. Projects carry out a variety of activities in their intervention areas and it is impossible to single out the impact of one specific intervention using quantitative methods. We therefore draw on the analysis presented in the projects' own endline evaluation reports to unpack how changes have come about and how effective different activities have been at creating these changes.

2.3 EM sampling approach

The EM sampling points were defined as the catchment areas surrounding each project's treatment and control schools, as per lists provided by projects at the start of the GEC. This sampling design was based on the assumption that the impact of interventions would have a sufficient magnitude to be picked up within local communities. Nevertheless, there is a possibility that some of the girls sampled attend schools that are further

away and as such are not receiving any GEC intervention. These caveats are explored and discussed in more detail in Box 1.

The projects' and EM's samples are independent from each other and are not representative of the same populations. While projects' treatment groups are supposed to be representative of the projects' target girls, the EM treatment groups are representative of the projects' target communities (school catchment areas), and the overall population of girls targeted by SCW projects as a whole (i.e. girls aged 5-15 at baseline). This explains why a substantial proportion of girls in the EM treatment groups do not attend GEC intervention schools and may not have directly benefitted from the GEC in-school activities. However, they should still have been exposed to community-based activities run by SCW projects in treatment areas.

In addition, girls in the EM sample might attend a treatment school but not fall within the specific age range targeted by the project, because the EM sampled all girls aged 5-15 at baseline in all project areas[45] for the purpose of consistency and comparability, even though several projects targeted only primary school girls (see Table 7). Therefore, **only project data should be used to quantitatively assess the impact and achievement of projects on their specific target populations**. **The EM sample has been designed to represent the SCW target communities as a whole and is not large enough to detect impact at the project level** (see Box 2). It is also important to remember that the age reported by girls and their parents may not always be accurate, because a lack of birth registration is common and families may estimate their children's birth date rather than know it for certain.

The EM endline research was conducted between October 2016 and April 2017 (see Section 1.4 of Annex E). It follows the EM baseline research that was conducted between May 2013 and June 2014, and the EM midline research that was conducted between August 2015 and June 2016. In each project area, the EM surveys roughly 400 households and conducts EGRA and EGMA tests with 400 girls that are being tracked longitudinally from baseline through to midline and endline. The sample of 400 households per project area is split evenly between treatment and control school catchment areas. See Annex E for further details on our sampling approach.

Box 1: Composition of the EM treatment and control samples in terms of GEC and non-GEC schools, and impact on the findings presented in this report

At midline, the EM conducted additional analysis to assert how many girls in the EM community sample attend SCW treatment schools, as opposed to other schools in the area. To this end, a list of all schools reportedly attended by girls in the EM sample was extracted from the midline data. Projects were then asked to identify which of these schools were targeted by their GEC interventions. Eight SCW projects provided this information.

Our analysis (described further in Annex F) showed varying patterns across project areas. A number of control groups contain large proportions of girls attending GEC schools. Similarly, some areas identified as treatment have less than one quarter of in-school girls enrolled in GEC intervention schools. Still, across all treatment areas, more girls are attending GEC intervention schools than in the corresponding control areas: on average, 61% of girls in treatment communities attend a GEC school, against 30% in control communities.

However, further analysis showed that at midline girls attending GEC schools had not improved significantly more in the GEC main outcomes (attendance, literacy and numeracy) than girls attending non-GEC schools[46]. These findings confirmed those found when using the EM treatment and control groups.

In this endline report, we therefore rely only on the comparison of treatment and control areas. Although there may be a mix of girls attending GEC and non-GEC schools in both types of areas, the purpose of our EM sample is to measure the effect of the SCW on communities rather than on target girls. Girls who do not attend a GEC school but who live in treatment areas (such as out-of-school girls) may still have been reached by GEC interventions run by SCW projects in their communities, which girls living in control areas may have not been exposed to. Besides, the analysis of our EM midline qualitative research showed that SCW treatment and control communities were crowded spaces where a range of education-related activities were operating. The

---

[45] In Camfed project areas the EM also sampled girls aged 16-17 at baseline, for reasons described in Annex E.
[46] Enrolment was not included in the analysis as the sample considered here is made up of in-school girls only.

comparison of GEC-school and non-GEC-school girls would have faced similar limitations. They are described further in Sections 2.5.1 and 3.1.

2.4 Summary of the SCW fieldwork processes

The SCW midline fieldwork was managed by ORB International and carried out by local research partners that were responsible for recruiting enumerators and supervisors and for overseeing the day-to-day fieldwork. In Annex E we provide a detailed discussion of our methodology and fieldwork processes that covers issues such as: sampling; interviewer training and piloting; fieldwork supervisions; non-response and attrition; and data processing.

2.5 Methodological limitations and mitigation strategies

2.5.1 Limitations of the EM endline research

Survey attrition

The EM household and school-visit survey is a longitudinal study which aims to interview the same girls at baseline, midline, and endline. The validity of estimates in longitudinal datasets depends on how successfully the study maintains respondents within the sample from one survey wave to another, to minimise any bias that can result from initial respondents dropping out of the study sample.

In the EM household survey, girls who could not be interviewed again at endline were replaced with girls from the same village (for a detailed description of the EM's replacement procedure see Section 1.3.1 and 1.5 of Annex E). Bias can be introduced when those who dropped out of the survey share specific characteristics, or if substitute respondents differ systematically from those who dropped out. Our analysis of attrition from the EM household survey does not show any significant difference between the cohort sample (girls recontacted from midline) and the full sample (including new girls interviewed at endline) in terms of key endline characteristics directly related to the GEC main outcomes and barriers (see Section 1.5 of Annex E).

Table 4 below shows attrition rates by project areas for the EM samples. Attrition rates have been mostly similar across treatment and control groups, except in CARE (Somalia) and World Vision (Zimbabwe). These differences are mostly due to entire sampling points being substituted rather than systematic losses to the sample in either the treatment or control groups.

Table 4: EM endline attrition rates by project area

Endline attrition EM HHS	All excl. Camfed	BRAC Afg	AKF Afg	ACTD Afg	IRC DRC	STC Eth	ChHp Eth	WUSC Ken	CfBT Ken	STC Moz	Plan Sie	RI Som	CARE Som	WV Zim	Camf Z-T
Treatment	26%	30%	25%	20%	29%	23%	9%	36%	33%	26%	26%	27%	22%	27%	36%
Control	25%	30%	21%	18%	33%	29%	8%	31%	37%	25%	23%	25%	32%	18%	
T-test (p-value)	0.83	0.97	0.31	0.66	0.35	0.22	0.68	0.24	0.47	0.95	0.48	0.60	0.02	0.03	

Notes: The third row shows the p-value of the statistical test of equality between two means (Student's test) comparing attrition rates across treatment and control areas. P-values lower than 0.05 (in red-filled cells) show a significant difference in attrition rates between the two groups at the 5% level.
Camfed's project areas are not associated with any control group in EM data. The reason is explained in the "Differences in EM research activities across different project areas" paragraph below and further in Annex E.

Limitations to the use of EM data at the project level

As per the GEC Evaluation Strategy, projects were intended to use a common data collection methodology (see Section 2.1.1) for their own research so that the EM could combine the window-level analysis of EM data with meta-analysis of project data. However, projects had to adapt the template methodology to the specific challenges they faced in their intervention contexts (see Section 2.1.1). As a result, not all projects are using control groups and not all projects are measuring learning outcomes in the same way or with similar age groups. In addition, projects struggled to merge their baseline, midline and endline datasets to the extent that we have not been able to use their quantitative data to measure quantitative changes in a rigorous way.

We therefore mostly relied on EM data to measure quantitative changes between baseline, midline and endline. The small size of the EM sample is a limitation, especially in project areas where attrition or refusal rates at midline were higher than anticipated (see Section 1.5 of Annex E). It is therefore possible that some changes that have occurred since baseline are not accurately captured in the EM sample at the project level (on top of the composition bias already discussed in Box 1). Box 2 presents a detailed discussion of the power of the EM sample.

Box 2: The statistical power of our EM sample

Levels of statistical significance are not only the reflection of the absolute size of a change but also depend on the sample size on which this change is being assessed. In our EM sample, project areas with higher attrition levels are therefore less likely to exhibit a significant change than other project areas. The same could be said for subgroups of smaller sample sizes, for instance secondary girls compared to primary girls.

Table 4 shows that attrition levels have been around 25 per cent at endline (compared to 30 per cent at midline) in our EM sample. A randomised controlled trial with a 30 per cent attrition per wave needs a minimum of about 1,000 individuals in the treatment group and the same number in the control group to be able to detect a change in outcomes of 0.2 standard deviations (SD) under the standard assumptions of a statistical power of 0.8 and a 95 per cent confidence level. Our EM sample has only 400 households per project area, which restrains our ability to detect small changes at the project level. It is however able to detect a 0.2 SD effect at the window level as its size is equal to about 6,000 households across all SCW project areas.

The EM samples are not designed to measure effects at a similar level of precision and statistical significance as the projects' larger datasets. They are designed to inform a robust impact evaluation at the window level among SCW target communities. Any finding expressed at a level of disaggregation lower than the window level (such as at the project level) must be interpreted with those caveats in mind.

Spill over effects and contamination

Quantitative and qualitative analysis shows that education programmes are widespread in SCW treatment and control areas. At midline, IDI respondents from EM research mentioned GEC activities taking place in areas assigned to the control group. It might be that treatment and control sampling points are too close to each other geographically to prevent spill over effects. Households may also have moved from one area to another since baseline. It is also possible that some GEC projects ended up working in schools or communities that they had originally designated to be part of control areas when they submitted their sampling lists to the EM in 2012. These limitations will be explored further in sub-section 2.5.2 and Section 3.1.

Where spill over effects have led to positive change in GEC outcomes or barriers in control areas we will struggle to identify a significant difference between treatment and control and to attribute a treatment effect to GEC activities. We therefore systematically compare and cross-check our findings with those presented by projects.

Differences in EM research activities across different project areas

No EM school-based research was conducted in Afghanistan to avoid security and safety issues arising for SCW projects and the EM's enumerators and research teams. For reasons related to Camfed's (Tanzania and Zimbabwe) project design (see below), no school-based research was conducted in their project areas.

At baseline, Camfed (Tanzania and Zimbabwe) did not provide a community-based listing because its intervention population was located exclusively within secondary schools. The EM therefore used a listing of the home communities of girls due to receive bursaries as a sampling frame. In these communities, a mix of randomly selected girls aged 5-15, like in other project areas, and of older girls aged 13-17, some of them having received or currently receiving a bursaries has been selected. See Section 1.3.1 of Annex E for more details.

Cultural and language differences across SCW contexts

Respondents' and interviewers' subjectivity are common issues in social research. They can be mitigated by enforcing rigorous quality controls (see Section 1.7 in Annex E) and triangulating data from different sources. The EM data collection covers a range of different contexts which differ widely in terms of social structure, livelihoods, attitudes and perceptions of education. For instance, a similar proportion of caregivers in Camfed (Tanzania and Zimbabwe) and Save the Children (Ethiopia) project areas feel there is not enough support in their communities for

girls to succeed in school. But further research showed that community attitudes towards education among pastoralist communities in the Afar region are negative compared with Zimbabwe and Tanzania.

A similar caveat applies to the interpretation of outcome results: SCW projects cover areas and communities, which have different first languages and languages of instruction. Children have been tested in the official language of instruction. It is the same language they speak at home in some areas, but not in others. Nevertheless, we score all children against a common literacy metric (i.e. the EGRA oral reading score in words-per-minute) without accounting for such differences. Table 5 shows the official languages of instruction in each SCW project area, in which the EM EGRA and EGMA assessments were given, and the proportion of girls who speak a different language at home.

Table 5: Proportion of girls whose language of instruction is different from the language spoken at home

EM Data	All	BRAC Afg	AKF Afg	ACTD Afg	IRC DRC	STC Eth	ChHp Eth	WUSC Ken	CfBT Ken	STC Moz	Plan Sie	Relief Som	CARE Som	WV Zim	Camf. Zim	Camf. Tan
% of girls	51%	17%	22%	54%	85%	75%	6%	90%	91%	93%	96%	14%	7%	46%	31%	38%
Official language of instruction (LOI)		Dari, Pashto	Dari, Pashto	Dari	French	Amharic	Amharic	English Kiswahili	English Kiswahili	Portuguese	English	English Somali	Somali English	English	English	English Swahili

2.5.2 Limitations of the SCW projects' endline research

Many SCW projects faced issues with attrition, tight data collection timelines, finding skilled enumerators, matching of cohort observations across baseline, midline and endline, and the comparability of treatment and control groups. This section gives an overview of some key challenges and mitigation strategies reported by projects in their endline reports.

As shown in Table 6, a few SCW projects experienced **high attrition rates at endline** and many of them faced challenges related to the **longitudinal tracking of cohort girls and the matching of observations across survey waves**. AKF and BRAC (Afghanistan), IRC (DRC) and CARE (Somalia) reported difficulties locating cohort girls because of inaccurate contact details, and girls leaving their households to migrate or marry. All three Afghan projects also reported difficulties identifying the right girls within their households as the names or ages recorded at previous survey waves were sometimes inaccurate. Tracking of out-of-school girls was particularly difficult, with BRAC drawing on community leaders to locate at least some of these girls.

In some cases, challenges to the longitudinal tracking were so severe that the projects **abandoned the tracking approach altogether**. Save the Children (Ethiopia) adopted a cross-sectional sampling design at midline when a severe drought caused large-scale migration from their project areas. Plan (Sierra Leone) suspended their midline research due to the Ebola outbreak and related school closures, and largely adopted a cross-sectional approach at endline. WUSC (Kenya) abandoned the cohort tracking in response to significant population movements in the Kakuma and Dadaab refugee camps where they were operating.

Table 6: Attrition rates reported by projects

Attrition rates by project	BRAC-C	BRAC-G	AKF	Acted	IRC	STC	ChHpe	WUSC	CfBT	STC	Plan	RI	CARE	WV	Camfd	Camfd
	Afg	Afg	Afg	Afg	DRC	Eth	Eth	Ken	Ken	Moz	Sie	Som	Som	Zim	Tan	Zim
BL-ML	12%	30%	65%	no ML	44%	N/A	15%	N/A	49%	10%	no ML	27%	50%	35%	18%	35%
ML-EL	7%	19%	12%	no ML	47%	N/A	8%	N/A	44%	19%	no ML	39%	30%	18%	6%	27%
BL-EL	18%	43%	69%	<5%	70%	N/A	22%	N/A	71%	27%	79%	55%	65%	47%	23%	53%

Note: The table shows attrition rates for the learning cohort only. Attrition of other surveys may differ from the rates shown above.
BRAC-C represent girls enrolled in BRAC Community schools, BRAC-G girls enrolled in BRAC Government schools. Because they form two separate cohorts of girls, attrition rates are shown separately, as for Camfed Tanzania and Camfed Zimbabwe.

Several projects discovered **errors in their baseline and midline data, or flaws in the matching of baseline and midline observations at endline**. CARE (Somalia), for instance, found duplicate entries in their existing datasets, which led to a smaller than expected final endline sample. Save the Children (Mozambique) realised that some individual identifiers had been wrongly recorded and that in some households more than one girl had been surveyed in a breach of the survey protocol, requiring them to re-clean the baseline and midline data.

Insecurity remained a major challenge for projects working in conflict-affected communities. Some sampling points in Baghlan and Badakhshan (AKF (Afghanistan)) and North and South Galkayo (RI (Somalia)) had to be replaced, and CARE (Somalia) had to substitute 10 out of 55 schools that had become inaccessible due to drought, insecurity, or fire. Insecurity in Dadaab and Kakuma refugee camps and host communities (e.g. kidnappings, terrorist attacks) caused school closures and absenteeism amongst students and teachers in WUSC (Kenya) project areas, which led to attrition from the survey and learning tests. Other projects did not lose any sampling points but had to adjust their fieldwork schedules and accommodate elevated levels of suspicion amongst research participants (e.g. BRAC and ACTED, Afghanistan).

Some projects observed significant differences between the **treatment and control groups** (e.g. ChildHope (Ethiopia), CFBT (Kenya), World Vision (Zimbabwe), IRC (DRC), and Save the Children (Ethiopia)). In addition, some projects found that the **girls and communities that had dropped out of the sample between midline and endline were significantly different from those remaining in the sample**, or those sampled for substitution (e.g. CARE Somalia, ChildHope Ethiopia, and IRC DRC). In both instances, such differences can confound the measurement of treatment effects at endline.

Some projects faced **challenges that related to their qualitative research**. AKF and BRAC (Afghanistan), CARE and RI (Somalia) struggled to recruit vulnerable groups (e.g. married girls, ethnic minority girls) for qualitative interviews and focus groups. AKF, ACTED (Afghanistan) and CARE found that transcripts were of varying quality, and CARE acknowledged that the qualitative analysis could have been more nuanced if done in the original language rather than based on translations. RI (Somalia) and WUSC (Kenya) acknowledged potential bias in the qualitative data due to researchers asking leading questions or project staff recruiting respondents.

The EM midline report suggested that GEC control communities may have benefitted from GEC or external interventions, leading to the so-called **'contamination' of the control group** for research purposes. At endline, World Vision (Zimbabwe) raised concerns about control girls transitioning to treatment schools, or to Camfed (Zimbabwe) secondary schools. It was likely that control schools had also received external interventions as the Zimbabwean government reportedly aims to spread NGO activities equally across the country.

Save the Children (Ethiopia) explicitly asked research participants to rank local interventions (GEC and non-GEC interventions) according to what had most affected their education and found that the World Food Programme (WFP) and UNICEF had a large presence in treatment and control areas, with the WFP's school feeding programme having a positive effect on education especially in control areas. Save the Children (Ethiopia) also suggested that their support to regional education authorities and other communication activities may have affected schools in control areas.

Some projects faced **challenges related to the relatively tight timelines of the endline fieldwork**, which happened only a year or less after the midline fieldwork. RI (Somalia) and Save the Children (Ethiopia) had to collect data at an earlier time in the school year cycle than at previous survey waves, so seasonal variations in livelihoods may have introduced some inconsistencies into the data. World Vision (Zimbabwe) lacked time to prepare their research instrument and protocols, which led to some technical glitches and disruption of the early fieldwork.

Other **logistical challenges** involved poor road conditions due to heavy rains (Plan (Sierra Leone), World Vision (Zimbabwe)), difficulties communicating with schools to arrange visits (Plan (Sierra Leone) and Relief International (Somalia)), technical challenges around connectivity in the field and difficulty uploading data and correcting scripting errors in real time (e.g. Plan (Sierra Leone), World Vision (Zimbabwe), Camfed (Tanzania and Zimbabwe)). In other cases, difficulties stemmed from the poor administration of the survey and tests. ACTED, for instance, had to reject 75 surveys out of 500 due to quality concerns (i.e. consistency, missing data, implementation too rapid).

Ceiling effects and flooring effects were observed on parts of the EGRA and/or EGMA tests for CARE (Somalia), Save the Children (Mozambique), Relief International (Somalia), and WUSC (Kenya). Finally, the **quality, availability and reliability of school register data** remained a significant challenge, but many projects have adapted and triangulated existing school data with their own spot checks and head counts at endline.

3 Key Findings

3.1 To what extent has the SCW reached marginalised girls?

3.1.1 Who are SCW projects targeting?

The GEC aims to "expand education opportunities to marginalised girls".[47] Its business case defines marginalised girls as "those girls of primary and secondary school age […] who have not been enrolled or have dropped out from school (whether living in slums, remote areas, ethnic minorities, girls with disabilities etc.) or are in danger of doing so."[48] Marginalisation was defined in terms of the education outcomes that DFID wanted the GEC to focus on: enrolment, attendance and learning (literacy and numeracy).

DFID chose not to prescribe the marginalisation factors that projects should focus on to help girls attend school and learn, because of a lack of evidence about what caused girls' marginalisation from education in each context. GEC applicants were encouraged to identify the girls with the greatest educational needs in a given context (e.g. within a country, region or community). It was the projects' responsibility to understand and articulate what factors caused these girls to be marginalised from education.

Each project responded to the challenge of finding the most marginalised girls differently. Therefore, the SCW target population includes a variety of school status, grade levels, and socioeconomic subgroups. Table 7 shows the share of different subgroups in each SCW project's target population, according to projects' endline evaluation reports[49]:

- **School status:** All projects are targeting girls enrolled in school. Six projects also target out-of-school girls (not enrolled in school at the start of the project), who either dropped out or have never been enrolled.

- **Grade level:** All SCW projects, apart from Camfed (Tanzania and Zimbabwe), are targeting girls in primary school, most of them in lower primary school. Only 4 out of 14 projects target girls in upper secondary school. Camfed targets exclusively secondary school girls.

- **Socioeconomic subgroups:** Ten projects explicitly considered poverty as a criterion for identifying their target girls, even though the communities targeted by the remaining four projects are also affected by poverty. Eight projects target girls with a disability, seven projects target orphans, four projects target girls who have been displaced, four projects explicitly target young mothers, four target girl labourers and three projects target girls from pastoralist communities.

Table 7: SCW projects' endline target population by school status and socioeconomic subgroup

Endline target groups Projects' endline reports	Projects targeting this group	BRAC Afg	AKF Afg	ACTD Afg	IRC DRC	STC Eth	ChHp Eth	WUSC Ken	CfBT Ken	STC Moz	Plan Sie	RI Som	CARE Som	WV Zim	Camf Z-T
Total girls targeted		50100	38199	na	109577	12479	16503	17046*	89437	32732	21060*	47236	10145	60967	172826
School status															
OOS – never attended[1]	6			na	1%		9%			17%		7%	10%		
OOS – dropped out[2]	6			na	2%		3%					6%	7%		
Lower primary	12	100%	77%	na	55%	100%	23%		39%	71%	na	39%	17%	58%	
Upper primary	12		16%	na	41%		47%	na	61%	12%	na	34%	41%	36%	
Lower secondary	10		3%				14%	na		0%	na	8%	16%	6%	100%
Upper secondary	4		3%				4%					6%	9%		
Socioeconomic subgroups															
Disabled girls	8	1%					3%		2%	5%	na	4%	2%	23%	
Orphaned girls	7	2%				na			14%	11%		9%	9%		33%
Pastoralist girls	3				100%				16%				7%		
Displaced girls	4						na		2%			10%	3%		

---

[47] https://www.gov.uk/girls-education-challenge#girls-education-challenge--the-portfolio-of-projects
[48] DFID (2012), Girls' Education Challenge, Business Case Version 4, June 2012, p. 30.
[49] This information was not available in the endline reports of ACTED (Afghanistan), Plan (Sierra-Leone) and WUSC (Kenya).

Endline target groups Projects' endline reports	Projects targeting this group	BRAC Afg	AKF Afg	ACTD Afg	IRC DRC	STC Eth	ChHp Eth	WUSC Ken	CfBT Ken	STC Moz	Plan Sie	RI Som	CARE Som	WV Zim	Camf Z-T
Slum-dwellers	1								21%						
Child labourers	4						75%		2%	9%			16%		
Girls in poor households	10	97%	100%	na			na		29%	74%		66%	63%	48%	67%
Disadv. caste/ethnic minority	5			na			na		2%				na	29%	
Affected by HIV/AIDS	1								2%						
Young mothers/expecting	4						6%		1%	2%	11%				
Street Children	2						4%		0%						
Other groups[3]	3						12%	na	9%						

Cells marked with "na" imply that the corresponding figure is not available, although the group is part of the project's primary target groups. Numbers of total girls targeted marked with an asterisk were extracted from the endline outcome spreadsheets.
1: "OOS girls" refers to out-of-school girls (that is girls who are not currently enrolled).
2: "Girls dropped out" refers to girls who were enrolled in the past but dropped out of school prematurely.
3: "Other groups" include: WUSC: Girls living in host and refugee communities; ChildHope: Girls affected by risky migration: children leaving their homes on their own (unaccompanied by parents/guardians) or accompanied by a trafficker, to areas where they will be vulnerable to physical, labour, economic and/or sexual abuse and exploitation; CfBT: Slum dwellers.

Several projects which report targeting orphans or disabled girls consider these characteristics when assessing girls' eligibility for bursaries but do not tailor a significant part of their interventions to specifically supporting these target groups. These subgroups therefore only account for a small share of their overall target population. This can be explained by the difficulty of identifying a large proportion of girls belonging to these subgroups within the projects' areas of intervention. When subgroups have been included in the projects' samples, their size is sometimes too low to provide meaningful disaggregation. In sub-section 3.3.6, we present an **analysis of outcomes by subgroups** based on projects' samples which includes a discussion of sample size limitations.

It is important to bear in mind that some projects sampled girls of various ages in their households while others sampled and tested girls in specific grades at school. Each project sample is representative of the groups that the project is targeting and therefore not exactly comparable to other SCW project samples.

Table 8 below shows the number of girls actually reached by each project (direct learning beneficiaries), as well as the number of other beneficiaries among target communities and schools (other girls, boys, adults and teachers). The number of other beneficiaries have been directly submitted by projects and may therefore be inaccurate. Based on available information, most SCW projects reached their target numbers of direct beneficiaries at endline. Plan (Sierra Leone) is an exception as the project reached only about one third of their targets. This is largely due to school closures resulting from the 2014 Ebola outbreak, which restricted Plan's (Sierra Leone) area of intervention across the country.

Table 8: Number of beneficiaries (reached) by SCW project

Total beneficiaries reached	BRAC Afg	AKF Afg	ACTD Afg	IRC DRC	STC Eth	ChHp Eth	WUSC Ken	CfBT Ken	STC Moz	Plan Sie	RI Som	CARE Som	WV Zim	Camf Z-T
Direct beneficiaries	50 100	38 199	15 024	109 577	12 479	17 773	17 046	88 921	45 423	7 362	47 236	28 865	45 859	171 640
In-school girls	50 100	38 199	4 124	105 521	12 479	16 134	17 046	88 921	37 864	5 079	47 236	23 039	45 859	171 640
Out-of-school girls	0	0	10 900	4 056	0	1 639	0	0	7 559	2 283	0	5 826	0	0
Other beneficiaries	164 103	358 069		195 692	36 931	17 917	238 431	286 284	131 069		251 146	18 641	288 556	385 732
Other learning beneficiaries. (boys)	0	18 022		21 806	9 190	?	26 693	89 822	32 000		57 474	8 709	0	184 223
Broader student beneficiaries. (boys)	0	99 062		102 117	9 190	3 035	58 897	89 822	39 000		70 184	0	60 000	96 076
Broader student beneficiaries. (girls)	100 000	112 225		42 649	9 190	6 240	39 975	89 437	45 423		7 008	2 536	101 448	87 814
Teacher beneficiaries	3 670	5 917		4 412	1 133	2 858	1 919	3 168	2 299		835	631	2 931	4 545
Broader community beneficiaries (adults)	60 433	122 843		24 708	8 228	5 784	110 947	14 035	12 347		115 645	6 765	124 177	13 074

3.1.2   How marginalised are the girls that are being targeted through the SCW?

The rigorous and systematic targeting of only the most marginalised girls is not a GEC programme requirement. While many of the girls targeted by GEC projects can be classified as marginalised, and even severely marginalised, there is no evidence that they are the most marginalised in their project areas.

Both the EM and projects sampled girls from target communities and comparable control communities, to assess the impact of GEC interventions on their education outcomes, and did not sample girls from outside these communities. Thus, we cannot assess how marginalised GEC target girls are in relation to girls from non-target communities. Still, we can compare the ways in which girls are marginalised across different SCW contexts. Table 9 shows a set of "markers" of marginalisation that indicate the severity of factors that may hamper girls' educational outcomes. These are variables from the EM endline household survey which were chosen because of their relatively high reliability and low non-response rates. In Table 10, we provide a comparison of some of these indicators with national-level secondary data.

- **Markers of "educational marginalisation"** include low enrolment (% of girls who had never attended school at baseline), poor literacy (% of girls that had an oral reading fluency score of zero at baseline), speaking a different first language than the language of instruction, and having a primary caregiver who is unable to read or write a letter in the language of instruction.

- **Markers of "economic marginalisation"** include the affordability of schooling, landlessness, inability to meet basic needs, and hunger (% of caregivers stating that a member of the family has gone to sleep at night feeling hungry more than 5 days in the past year).

- **Markers of "social marginalisation"** include the absence of parents, disability, and safety (% of caregivers saying it is fairly or very unsafe for girls to travel to schools in the area).

Levels of marginalisation are reported by household members themselves and are therefore subjective. They need to be understood within each project's cultural and institutional context. For instance, although poverty levels are relatively high across the SCW, the difficulties that survey respondents report in relation to affording school largely depend on whether households must pay for school fees and materials, which in turn depends on government policies and/or school practices at the local level as well as on the age and grade of children (secondary school fees being usually higher than primary school fees).

Table 9: Markers of marginalisation by project area

Markers of marginalisation EM endline data	All	BRAC Afg	AKF Afg	ACTD Afg	IRC DRC	STC Eth	ChHp Eth	WUSC Ken	CfBT Ken	STC Moz	Plan Sie	RI Som	CARE Som	WV Zim	Camf Zim	Camf Tan
Educational Marginalisation																
Poor Attendance - Girl had never attended school at baseline	18%	22%	17%	19%	10%	32%	16%	38%	5%	4%	11%	39%	43%	2%	4%	9%
Poor Literacy - Girl scored zero wpm at baseline	54%	48%	45%	48%	64%	96%	42%	78%	44%	56%	62%	42%	46%	38%	33%	36%
Language difficulties - LOI different from language spoken at home	53%	11%	5%	56%	82%	73%	11%	91%	90%	91%	92%	22%	13%	53%	32%	16%
Parental literacy - PCG cannot read or write a letter in LOI	71%	90%	94%	97%	40%	99%	82%	93%	57%	51%	76%	53%	41%	57%	55%	32%
Economic Marginalisation																
Affordability of school - Difficult to afford for girl to go to school	51%	33%	7%	17%	83%	10%	24%	67%	89%	39%	81%	62%	66%	80%	88%	56%
Land ownership - Household doesn't own land for themselves	30%	43%	31%	22%	32%	19%	11%	50%	68%	9%	39%	64%	25%	12%	19%	17%
Poverty - Household unable to meet basic needs	41%	32%	21%	44%	44%	41%	30%	50%	46%	28%	40%	38%	43%	60%	69%	36%
Hunger - Gone to sleep hungry more than 5 d. in past year	14%	1%	1%	5%	13%	15%	4%	21%	21%	12%	62%	5%	15%	11%	16%	12%
Social Marginalisation																
Absent parents - Girl lives without her father and mother	7%	0%	0%	0%	9%	4%	6%	7%	8%	2%	14%	9%	5%	37%	31%	18%

Markers of marginalisation EM endline data	All	BRAC	AKF	ACTD	IRC	STC	ChHp	WUSC	CfBT	STC	Plan	RI	CARE	WV	Camf	Camf
		Afg	Afg	Afg	DRC	Eth	Eth	Ken	Ken	Moz	Sie	Som	Som	Zim	Zim	Tan
Disability - Girl has disabilities	9%	0%	1%	0%	15%	1%	8%	12%	17%	13%	17%	7%	7%	10%	14%	9%
Safety - Fairly or very unsafe travel to schools in the area	13%	6%	6%	7%	12%	32%	18%	3%	11%	10%	12%	2%	5%	43%	16%	9%

Table 9 shows that girls face different degrees of educational, economic, and social marginalisation across SCW project areas. Darker shades of red indicate higher prevalence of the related marker of marginalisation.

- There is a particularly strong concentration of educational disadvantage in WUSC (Kenya) and Save the Children (Ethiopia) project areas. As mentioned above, both projects work in particularly challenging contexts, with WUSC operating in Kakuma and Dadaab refugee camps and the local host communities, and Save the Children working with pastoralist communities in Ethiopia's remote Afar region.

- In all three project areas in Afghanistan the large majority of girls (over 90%) have a primary caregiver (PCG) who is illiterate in the language of instruction, with ACTED and AKF project areas showing higher illiteracy rates than any other SCW project area. Most of them being women, PCG's literacy rates may be a better indicator of women's status in certain contexts than of educational marginalisation. On all other indicators, however, the strongest rates of marginalisation are measured in African project contexts.

- Camfed and World Vision project areas in Tanzania and Zimbabwe show almost an inverse pattern to other SCW contexts in that educational marginalisation is relatively low, but economic and social markers of marginalisation show higher prevalence than elsewhere. In both project areas, the percentage of children who live without their parents is much higher than in any other SCW project area. Our midline qualitative interviews and thematic research in Zimbabwe indicated that adults sometimes migrate to South Africa or Botswana for work and send back remittances to help their children through school. Living without parents may therefore denote an educational advantage rather than a disadvantage[50].

Table 10: Markers of marginalisation – comparison with national-level secondary data

Secondary data	Afghan.	DRC	Ethiopia	Kenya	Mozamb.	Sierra-Leone	Somalia	Zimbabwe	Tanzania
Educational Marginalisation									
Girls' Primary Net Attendance Rate[51]	48% (2011-12)	85% (2013)	67% (2014)	88% (2008-9)	76% (2011)	78% (2013)	19% (2006)	94% (2014)	83% (2010)
Girls' Primary Completion Rate[52]	No data	60% (2013)	53% (2014)	104% (2014)	45% (2014)	68% (2013)	No data	90% (2013)	77% (2013)
Female aged 15+ literacy rate (2015)[53]	24%	67%	41%	75%	45%	38%	No data	85%	73%
Economic Marginalisation									
Population in multidimensional poverty[54]	59% (2011)	73% (2014)	88% (2011)	36% (2014)	70% (2011)	78% (2013)	82% (2006)	29% (2014)	66% (2010)

Table 10 shows statistics at the national level for each of the SCW target countries. It is worth noting that the years for the national statistics are not consistent across the countries, and the data for Somalia and Kenya lag significantly behind (2006 and 2008-9 respectively), while the EM data from Table 9 was collected in 2015/2016.

Secondary data confirms that some SCW project areas fall substantially behind national averages. For instance, WUSC's target communities (refugee camps and local host communities) are more marginalised than the average Kenyan girl, both from economic (multidimensional poverty compared to households unable to meet basic needs) and educational (primary net attendance rate compared to average attendance from EM survey data) perspectives.

---

[50] Evaluation Manager Girls' Education Challenge Fund (2016). Narrow windows and revolving doors. GEC Thematic Research Report, January.

[51] http://data.unicef.org/topic/education/overview/ Retrieved on 27 January 2017. Primary Net Attendance Rate as defined by UNICEF: Number of children attending primary or secondary school who are of official primary school age, expressed as a percentage of the total number of children of official primary school age.

[52] http://data.worldbank.org/indicator/SE.PRM.CMPT.FE.ZS

[53] https://www.cia.gov/library/publications/the-world-factbook/fields/2103.html

[54] http://hdr.undp.org/en/indicators/38606

Within the same country, CfBT's target group is closer to the national average in terms of poverty prevalence and enrolment levels. In Zimbabwe, World Vision and Camfed target communities also appear more economically marginalised than the national average, although the indicators used in the available secondary data and EM data may not be directly comparable.

Increased marginalisation due to recent external factors

Some SCW projects work in very challenging environments, having to accommodate conflict and instability, environmental disruptions, and disease.

In several areas of Afghanistan, insecurity and instability have increased over the past two years. For example, Kunduz and Baghlan province saw a rise in threats and attacks by anti-government groups, leading to the displacement of communities and temporary school closures. In 2016, close to 200,000 people in the North-Eastern region were displaced by conflict, with Kunduz being the most affected province (BRAC Endline Evaluation Report). Both AKF and BRAC have project communities in conflict-affected areas of Afghanistan.

In Mozambique, tensions rose between government and anti-government forces in 2015 which led to rebel attacks in Tete Province and Manica provinces. Check points were erected and mobility within Mozambique was very limited, making many schools in some of Save the Children's target areas inaccessible for most of 2015/2016.

Several project areas struggled with environmental disruptions and severe drought in 2015 and 2016, which have affected school attendance and learning. According to Child Hope's endline report (Ethiopia), the drought caused school closures in several regions of Ethiopia, and a lack of water and school feeding activities in remaining schools led to increased student absenteeism. At the same time, girls started to report higher working hours, suggesting an increased pressure to support their household through paid or domestic work. This is likely to have caused further absence from school.

In Zimbabwe, World Vision reported that the drought had caused rural families to struggle to grow and sell enough food to sustain themselves and pay for their children's education. Rising food prices made many families unable to provide three meals a day, so that children had to go to school hungry, which may have hindered their learning. In Somalia, over 6 million people were facing acute food insecurity in December 2016. The drought led to the loss of livestock and forced families to migrate, and some of Relief International's target schools were reporting decreasing student numbers. As of February 2017, four schools targeted by Relief International in Somaliland were closed due to drought with no immediate plans to reopen.

3.1.3   To what extent have communities been exposed to GEC interventions?

We use EM data to measure the extent to which target communities have been exposed to GEC interventions. We do not report on projects' own data because most questions related to exposure and reach were not systematically included in projects' surveys. When they were, endline data could not always be linked with data from previous waves. Out of 14 SCW projects, only 5 projects submitted consistent quantitative exposure and reach data across midline and endline, and only three did so for both treatment and control areas. Graphs based on our reanalysis of projects' data are shown in Annex K.

The EM surveys (like projects' surveys) ask respondents about activities focused on girls' education without explicitly mentioning the GEC. It is therefore possible that some of the changes observed in the household and school visit surveys are not directly due to GEC interventions. Conversely, respondents may not be aware of all the activities supported by the GEC that are ongoing, and of the organisations running them.

Besides, survey-based measures of exposure and reach are **subject to recall and subjectivity biases**. Reported measures, whether from household members (here the girl's primary caregiver) or school personnel (the girl's school administrator or head teacher), do not necessarily reflect the type and scope of the activities delivered in the field: respondents are more likely to reference tangible activities that they have directly received, seen or heard of.

Finally, it is worth reiterating that **the EM sample was designed to be representative of target communities, not of projects' direct beneficiaries**. This implies that the EM's analysis cannot be

used to draw conclusions about the extent to which projects' activities effectively reached their target populations.

Box 3: SCW project areas are crowded spaces: implications on the ability of samples to detect changes

At midline, we showed, based on evidence from the projects' reports and from the EM qualitative research, that SCW project areas are crowded spaces for education interventions. **Non-GEC education-related activities are taking place in several projects' target treatment and control areas**, often provided by international NGOs and relief agencies. Besides, **GEC projects themselves sometimes ended up operating in communities they initially defined as control** (see Section 2 and Annex F for more details). In other cases, spill over effects resulted from control communities mobilising themselves to replicate an activity carried out in a nearby GEC treatment area.

The resulting biases in the EM and projects' samples reduces our ability to detect changes and to attribute them to GEC activities. The absence of significant differences across treatment and control areas does not necessarily imply that the GEC activities have not had an impact on their target communities. The following sections on exposure and impact must be interpreted with these caveats in mind.

!

In this section, as in the rest of Section 3, we report difference-in-difference indicators from our EM data in the following way:
- marked with a **single asterisk** when significantly different from zero at the **10% level** (p-value < 0.1);
- marked with **two asterisks** when significantly different from zero at the **5% level** (p-value < 0.05).

As in the rest of the report, DID indicators of proportions (%) are reported in percentage points (pp).

Figure 2 shows the proportion of primary caregivers who reported specific activities across all fourteen SCW treatment areas. These questions were asked "since baseline" at midline and endline (and "over the past year" at baseline) which implies that reported proportions should increase or stay constant between midline and endline (as opposed to questions that are asked "since last year"). A null or small difference between two waves suggests that primary caregivers haven't been aware of any activity taking place during that time interval.

At endline, none of the activities mentioned in Figure 2 have been reported by more than 18% of primary caregivers in SCW treatment areas. Proportions are the same in control areas. However, some activities may be underreported because primary caregivers may not be aware of girls benefitting from an intervention, especially if it takes place at school.

Indeed, the same questions asked to school teachers show slightly higher prevalence: at endline, 30% of them reported that the girl has attended a special class or study group in the past year, and 27% of teachers reported that the girl received special tutoring or help with her school work. Only 9% of them reported that the girl had received a scholarship or bursary, a proportion similar to caregivers (10%). This may be explained by the fact that scholarships can be distributed either through schools or households.

Figure 2: Household survey exposure (PCG) across SCW treatment areas
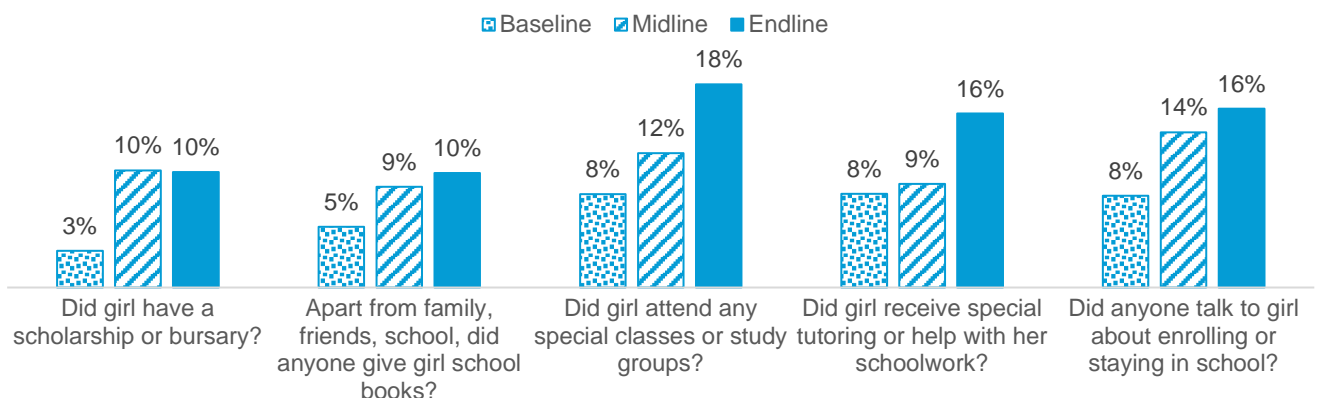
The school-visit survey contains a set of questions to school administrators asking about measures or activities to support girls' education taking place in their school. The first variable asked about any additional funds provided to the school over the past year. The second variable from the left asked about any activities taking place "over the past year" at baseline; and "since baseline" at midline and endline. The two variables on the right report on any measures that were still ongoing at the time of the survey. Figure 3 shows that school administrators reported an overall increase in all activities between baseline and endline, while the provision of additional funds kept constant.

Figure 3: School-administrator survey exposure across SCW treatment areas
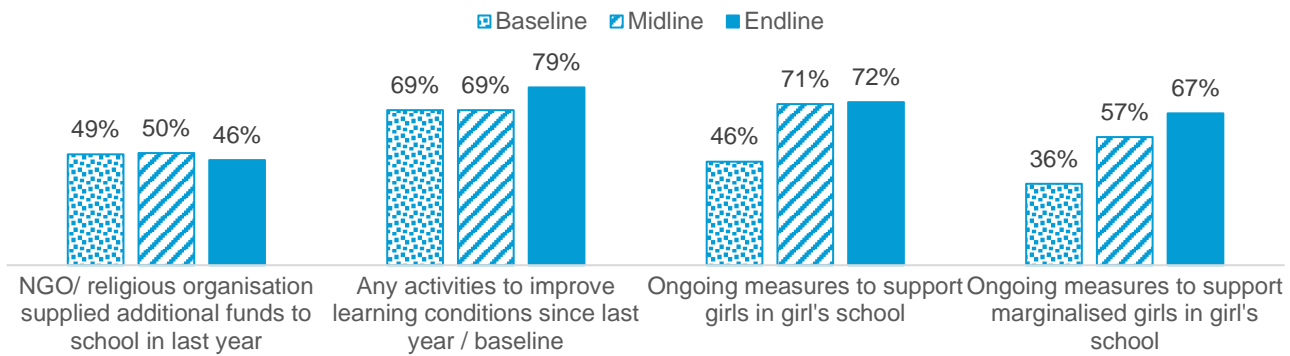

The reporting of activities to improve learning conditions have remained almost stable since baseline, which may be due to this question being broad and capturing any intervention aimed at improving learning, whether it has been implemented by the government, local authorities, the school itself or by non-governmental organisations. Conversely, the proportion of school administrators reporting ongoing measures to support girls has surged from 46% at baseline to 72% at endline. Measures supporting marginalised girls specifically have seen an even steeper increase from 36% to 67% at endline. In the two cases, most of the improvement has taken place between baseline and midline.

The proportion of administrators stating that an NGO or religious organisation has supplied additional funds to the school in the last year has remained stable since baseline. **However, it is the only activity that has been reported significantly more than in control areas between baseline and endline** (+19 pp*). This suggests that the provision of additional funding to schools is significantly more widespread in GEC areas. All other activities have been reported in equal proportions across treatment and control groups. This confirms the midline finding that a lot of activities related to girls' education have taken place in SCW treatment and control communities since baseline, some of which were driven by non-GEC actors, as stated in Box 3.

Figure 4 shows the proportion of school administrators reporting participation of girls' teachers in different types of training over the last year. These variables were added to the school visit survey at midline. This implies that they might have increased since baseline, like some of the variables described earlier, but our survey cannot capture this change. Between midline and endline, teachers' participation in trainings exhibit almost no variation. At endline, general teaching methods is the most commonly reported training (72%), before gender sensitive pedagogy (58%), mentoring or coaching (51%) and special needs education (28%). Again, none of them shows a statistically significant DID between baseline and endline at the window level.

Figure 4: Last year, teachers participated in training in...
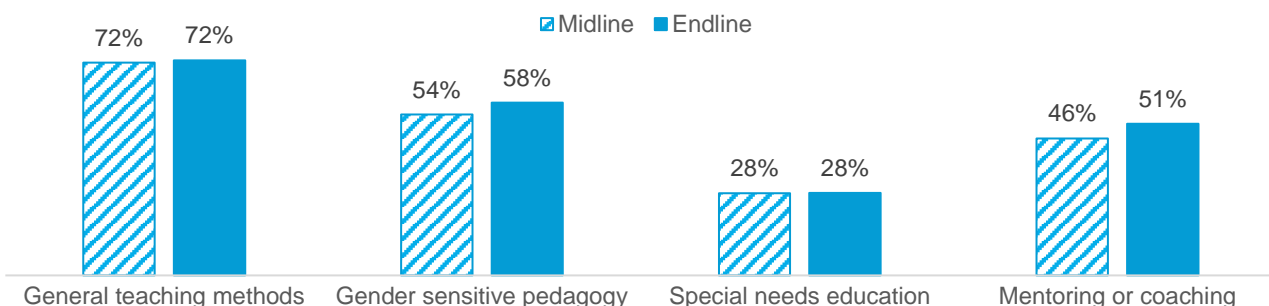
3.1.4 Who has been exposed to GEC interventions?

Using EM data, we explore the exposure to education-related activities of specific subgroups of girls: age categories, schooling status (never enrolled, dropped out, lower primary[55], upper primary[56], secondary, and girls with poor attendance) as well as personal and socioeconomic status: girl's parents not members of the household, primary caregiver cannot read or write a letter in language of instruction, girl does not speak the official language of instruction, girl has a physical disability, schooling difficult to afford, and household unable to meet basic needs.

It is worth noting that orphaned girls have been targeted by as many SCW projects as disabled girls (eight projects) but do not appear in our data. Our household survey does not ask whether the girl is an orphan, but whether their parents are members of the household. This subgroup therefore encompasses girls whose parents have left to work outside the country, which is a common situation in rural areas of Zimbabwe for example, especially those targeted by World Vision and Camfed (see Section 3.1.2). Our EM sample may just contain a few "true" orphaned girls across the SCW, which makes it difficult to draw a firm conclusion about whether they have been effectively reached by projects.

> **!** Due to sample size limitations, we look at subgroups across all SCW project areas at the same time. Thus, the absence of an observed effect on a specific subgroup could mean that some projects did not focus their activities on this subgroup. Other projects may have done so, but their effect might be diluted across the entire window.

The provision of scholarships or bursaries has been significantly higher for girls for whom schooling is difficult to afford (+3 pp**) which **suggests that SCW projects have been successful in providing scholarships or bursaries to households in need**. However, the difference is small in absolute terms and suggests that many households who declared that schooling is not difficult to afford did receive scholarships or bursaries.

The **provision of school books; special classes of study groups; special tutoring and help with school work; and talks about enrolling or staying in school**, all show the same pattern of change between baseline and endline: they **have reached enrolled girls significantly more (+6 pp**, +4 pp**, +3 pp** and +3 pp*) than out-of-school girls**. Indeed, these activities have often been provided through schools rather than through households.

The **provision of school books and talks about enrolling or staying in school has specifically reached girls with poor attendance in school** (+3 pp* and +6 pp** respectively). This suggests that those girls have been targeted specifically among all enrolled girls, which is relevant considering they may be at higher risk of drop-out.

Special classes, study groups as well as special tutoring or help with school work have been provided to upper primary girls significantly more than to other grades (+4 pp** and +3 pp** respectively): grades five to eight have been the most targeted by SCW projects for the running of in-school activities. Conversely, girls enrolled in secondary schools have received talks about enrolling or staying in school (+4 pp), although this increase is not statistically significant. These girls are often subject to higher risks of drop-out because secondary school can be more expensive, and as they grow older can experience early marriage, pregnancy or having to work inside or outside the household. Our finding suggests that SCW projects were conscious of those risks and tried to mitigate them by targeting secondary girls specifically.

Finally, it is worth noting that **girls with physical disability received significantly more talks about enrolling or staying in school (+12 pp**) and more special tutoring or help with school work (+6 pp*) than other girls**. They are the only socioeconomic subgroup standing out in our data as having been reached more often by specific projects' activities (except girls whose schooling is difficult to afford who have been reached by bursaries more, as mentioned above). This finding confirms the targeting of SCW projects shown in Table 7, which shows that disabled girls are the second most targeted social group behind girls living in poor households, and suggests that disabled girls may have been reached successfully.

---

[55] Defined as girls enrolled in Primary 1 to Primary 4.
[56] Defined as girls enrolled in primary school, in any grade higher than Primary 4.

Key findings – To what extent has the GEC reached marginalised girls?

Target and reach

- Projects targeted girls from a variety of schooling levels. Lower and upper primary school girls have been targeted the most (12 projects out of 14), followed by secondary school girls (10 projects) and out-of-school girls (6 projects).

- In terms of socioeconomic subgroups, projects targeted girls in poor households most frequently (10 projects), followed by disabled girls (8 projects) and orphaned girls (7 projects).

- Most SCW projects reached their target number of beneficiaries at endline. The number of girl beneficiaries ranges from less than 10,000 girls (Plan (Sierra Leone)) to about 170,000 girls (Camfed (Tanzania-Zimbabwe)). Unfortunately, the number of girls reached by schooling status and socioeconomic subgroup has often not been reported, and it was not a GEC programme requirement to track their learning outcomes specifically.

- While we cannot ascertain whether communities targeted by SCW projects are the most marginalised girls in their communities or regions, our data shows that **girls are disadvantaged across a variety of educational, social and economic factors whose prevalence and severity differ across project areas**. This reflects DFID's approach to allowing projects to define marginalisation as they see fit, and to identify contexts in which girls are marginalised from education.

Exposure

- **Levels of exposure to education support between baseline and endline were similar in treatment and control** areas across our EM sample.

- In treatment areas, **exposure to education-related activities has overall increased from baseline to endline**. **However, levels of exposure remain relatively low**: no activity is reported by more than 20% of caregivers as being currently received or having been received by girls since baseline. Note that we report averages across all SCW project areas, which may be subject to a "dilution effect" because activities have often been run only by a subset of projects.

- Low levels of exposure may also be due to a lack of awareness of caregivers, or difficulty to recall, especially about activities that usually take place in schools (such as the distribution of learning materials). Indeed, school surveys show slightly higher exposure levels, with some activities being reported by about 30% of school administrators (compared to less than 20% of caregivers). But there may also be incentives for school administrators to overstate the activity levels in their schools.

- At the subgroup level, **exposure variables show encouraging patterns that suggest that projects were, at least in part, able to reach the girls who are the most in need**: girls for whom schooling is difficult to afford have received significantly more bursaries or scholarships than other girls; girls with poor attendance and girls enrolled in secondary school have been talked to significantly more often about enrolling or staying in school; and girls with a disability have received significantly more tutoring or help with their school work.

Key lessons learned

- Although some projects targeted specific subgroups of marginalised girls, the extent to which subgroups were reached, and the extent to which their learning has improved as the result of the programme, remain unclear. Indeed, none of the subgroups were tracked as a specific cohort. **To assess the effect of GEC on subgroups of interest, they would ideally need to be represented in the projects' samples with sufficient sizes, and their outcomes tracked throughout the course of the programme**. This however would require substantial additional budget to be done consistently across GEC projects.

- The levels and dimensions of marginalisation differ across contexts. Therefore, activities aiming to improve girls' learning have to be context-specific and tailored to girls' needs. This highlights the

importance of **analysing contextual factors as an integral part of the design and M&E processes** – in particular, how and why different factors impact on the education of marginalised girls and the extent to which they constrain their capacity to stay in school and learn effectively.

- Our midline research showed that SCW communities are a crowded space for education programming, and that **GEC and non-GEC activities are taking place in most treatment and control areas, addressing similar educational barriers**. Respondents in control areas often mentioned having seen or heard of GEC project activities that were ongoing in neighbouring project target communities. This may be a positive sign of the SCW projects extending their impact (or at least awareness of their activities) beyond their intervention communities. But it may equally point towards inaccurate sampling frameworks provided by projects at the outset of the GEC. In both cases, it underscores the **difficulty of implementing valid quasi-experimental evaluation designs in the GEC context where actors, target populations and intervention designs keep shifting**. These "spill over" and "contamination" effects make it more difficult to statistically capture the effect of the GEC, both in terms of reach (as captured through exposure questions) and impact (on girls going to school and learning).

3.2 What impact has the SCW had on enabling marginalised girls to be in school?

As with the previous section, Section 3.2 and the subsequent Section 3.3 are mostly informed by the EM household and school survey data, except for the last two sub-sections 3.3.5 and 3.3.6.

Camfed project areas do not have a proper control group in the EM survey data. Difference-in-difference indicators for Camfed areas show the difference between two target communities rather than the difference between a treated and a non-treated group. To avoid any confusion when comparing treatment with control areas, aggregate effects across SCW exclude Camfed in the rest of this report. See sub-section 1.1.1 of Annex E for more details on the EM data methodology.

3.2.1 What effects has the SCW had on enrolment in school?

The enrolment rate increased from 76% to 82% since the start of the programme. This is less than +1 percentage point (pp) between midline and endline across EM SCW treatment areas, compared to a +4 pp increase between baseline and midline. As shown in Figure 5, enrolment increased to the same extent in the control group: from 75% at baseline to 82% at endline. Similarly, none of the SCW project areas shows a significantly higher increase in treatment than in control areas.

Figure 5: Enrolment rates across SCW project areas (excluding Camfed)
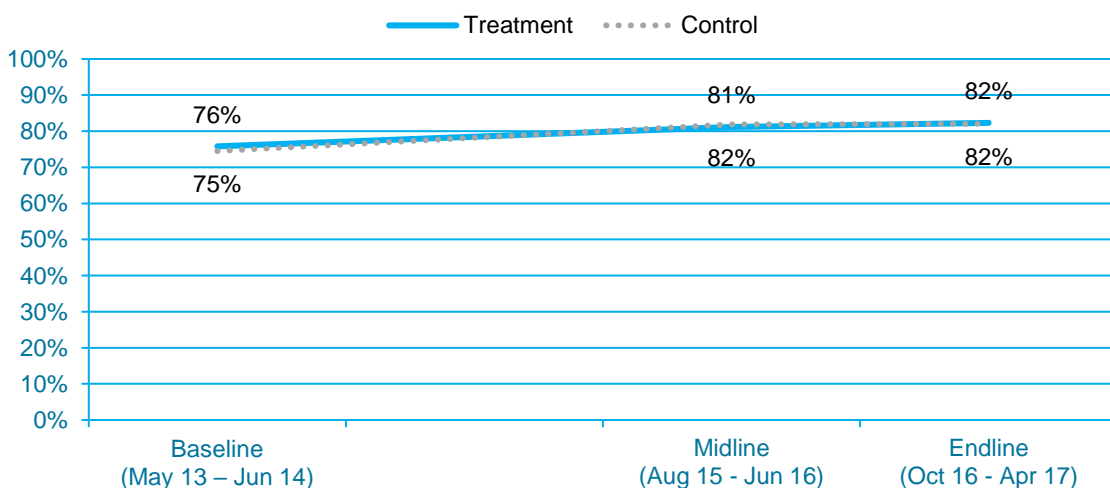
Aggregate enrolment rates hide **substantial variations at the project area level**. While in most project areas, more than 85% of girls are enrolled in school at endline, this rate is only 76% in Save the Children's (Ethiopia) and CARE's (Somalia) project target areas, 69% in WUSC's (Kenya) target areas and 65% in Relief International's (Somalia) target areas. When taking a closer look, the population of out-of-school girls is very different from one context to another.

Figure 6 shows the median number of years enrolled in school for in-school and out-of-school girls across SCW project areas at endline. It is worth noting that the out-of-school girls' sample sizes are very small for some projects (less than 50 cases) due to the small number of non-enrolled girls and the substantial proportion of girls for which the number of years enrolled cannot be calculated due to missing information about the age at which they started or stopped school. Although we can interpret large variations across contexts, these figures must not be taken as representative of the out-of-girl population within each project area.

Figure 6: Number of years enrolled in school at endline (median)
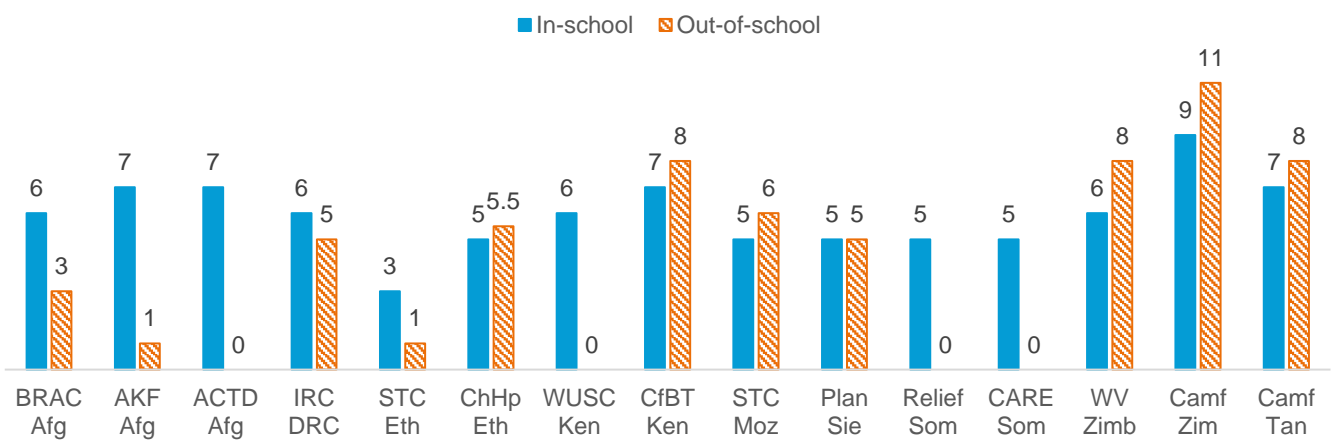

In about half of the SCW project areas, out-of-school girls have spent more years or about the same number of years in school as in-school girls. This suggests that a large proportion of out-of-school girls dropped out after having completed a cycle of education, in most cases primary school. It is worth noting that the median age of out-of-school girls varies across contexts with the lowest age in IRC's (DRC) target area at 13.5 years old and Save the Children (Mozambique) at 14 years old, and oldest in Camfed (Zimbabwe and Tanzania) at 17 years old, and Plan (Sierra Leone) at 18 years old.

In Afghanistan and Somalia, as well as among Save the Children's (Ethiopia) pastoralist communities and WUSC's (Kenya) target groups in refugee camps, the number of years that out-of-school girls have been enrolled in school is very low, indicating that these **girls have remained mostly outside the schooling system**. They are also younger, ranging from a median of 11 years old in CARE's (Somalia) target areas to 15 years old in BRAC's (Afghanistan) areas. The challenge in these contexts is therefore not so much about preventing drop-out from school, but more about enrolling girls or creating learning environments that can complement or substitute for school.

Figure 7 shows that across SCW, **girls are still lagging behind in terms of the school phase they are in.** At endline, more than 80% of girls aged 12-13 and more than half of the girls aged 14-15 are still enrolled in primary school, while almost all of them should be attending secondary school[57]. When they reach age 16-17, about half of the girls are enrolled in secondary school, but one quarter of them are out-of-school. As they grow older, drop-out increases: after the age of 17, almost half of them are out-of-school and only a minority have completed secondary school.

---

[57] It is worth noting that out of the nine SCW countries, seven (all except Afghanistan and Sierra-Leone) require the passing of an end of primary exam to progress to secondary school. Source: https://www.classbase.com/Countries and national sources.

Figure 7: School phase by age category at endline
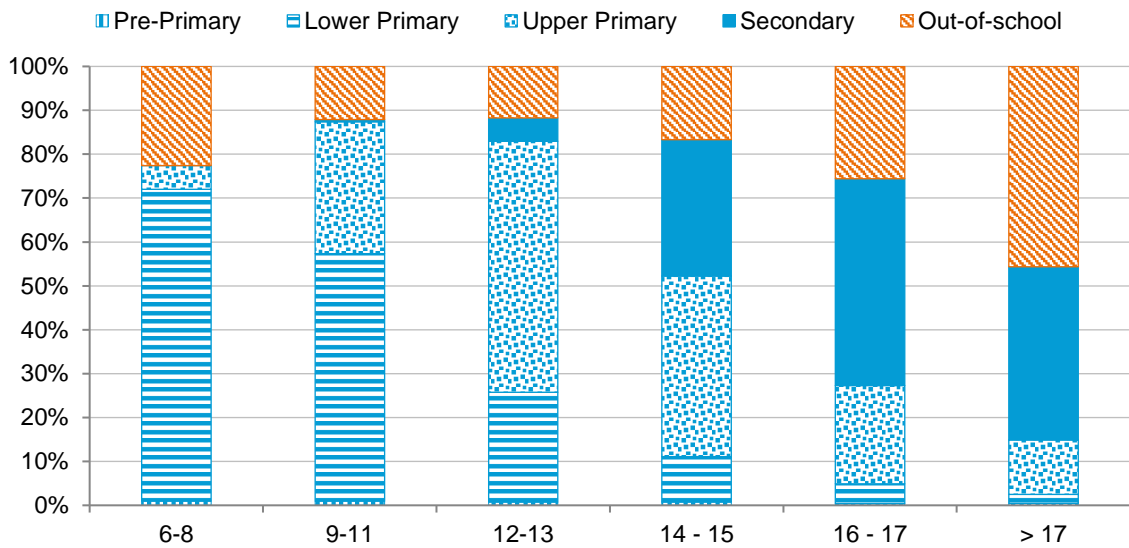
The above findings show that addressing girls' drop-out from school as they grow older still remains a challenge in most project areas. However, EM data suggests that SCW projects have been able to slightly slow down the drop-out of older girls compared to control areas. Figure 8 shows the percentage of girls enrolled at baseline who are still enrolled at endline. Although proportions are logically lower for older girls, we see that 79% of girls aged 14-15 at baseline in treatment areas are still enrolled in school, which is significantly higher than the 71% proportion observed in control areas.

Figure 8: Proportion of girls enrolled at baseline who are still enrolled at endline, by age at baseline
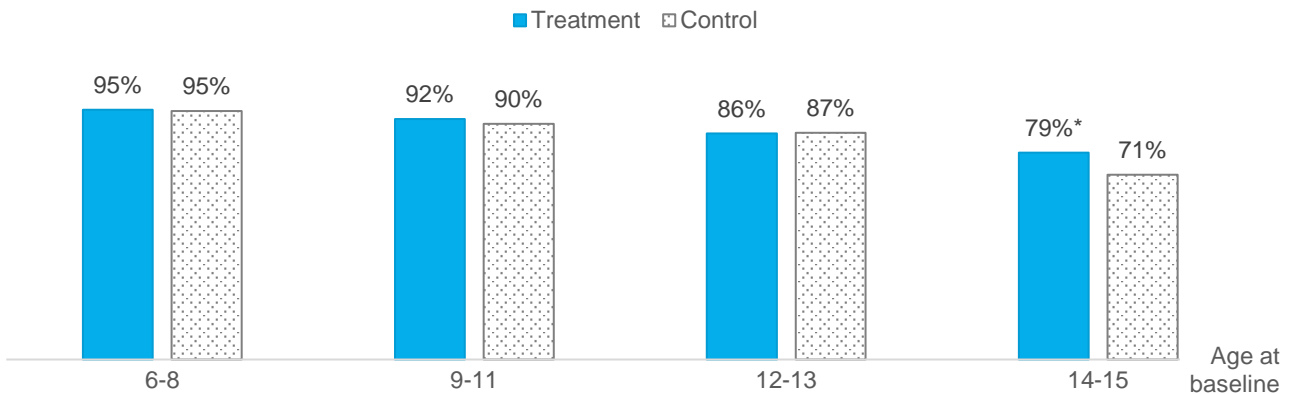
Conversely, the proportion of girls who were not enrolled in school at baseline and who are now enrolled is similar across treatment (65%) and control areas (63%). Due to low sample sizes, we cannot disaggregate this figure by age categories, but it is likely that most of this new enrolment is due to the youngest girls in the sample enrolling in pre-primary or primary school for the first time, which would imply that a large proportion of older out-of-school girls have remained out-of-school.

Finally, it is worth noting that boys' enrolment is now lower than girls', across all age categories, as shown in Figure 9. Within each household, one boy was randomly selected within the same age range as the girl (8-18 at endline) and questions asked to the primary caregiver about his enrolment in school, to be compared with girl's findings. Girls' enrolment rates increased overall, except for girls aged 9-11 where it was already high and remained stable, while boys' enrolment decreased across all categories. At endline, the gap now ranges from 3 pp to 6 pp in favour of girls. This pattern however is similar across treatment and control areas, which suggests that the observed trends may not be due to GEC activities.

Figure 9: Proportion of girls and boys enrolled in school at endline

■ Girls (Baseline) ■ Girls (Endline) ▨ Boys (Baseline) ■ Boys (Endline)
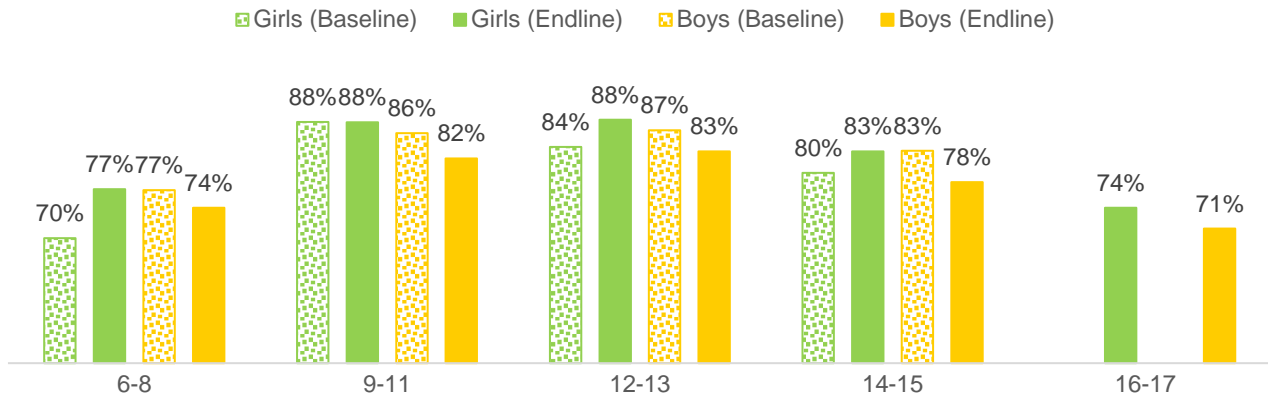
3.2.2 What effects has the GEC had on attendance in school?

Attendance is defined as the proportion of time girls are present in school when it is open. We use two distinct measures of attendance in our EM surveys: first, the household survey attendance rate, which is extrapolated from categorical questions asked to the primary caregiver. It is our core measure of attendance, but it is approximate and insensitive to small changes; and second, a measure based on the number of days a girl has been present (or absent) from school since the start of the school year, according to her head teacher. It is available only in our midline and endline data, in project areas where the school visit survey data was conducted (all except Afghanistan, Plan (Sierra Leone) and Camfed (Zimbabwe-Tanzania) project areas). Figure 10 shows the trajectory of attendance rates from our EM household survey measure.

Figure 10: Attendance rates across SCW project areas (excluding Camfed)
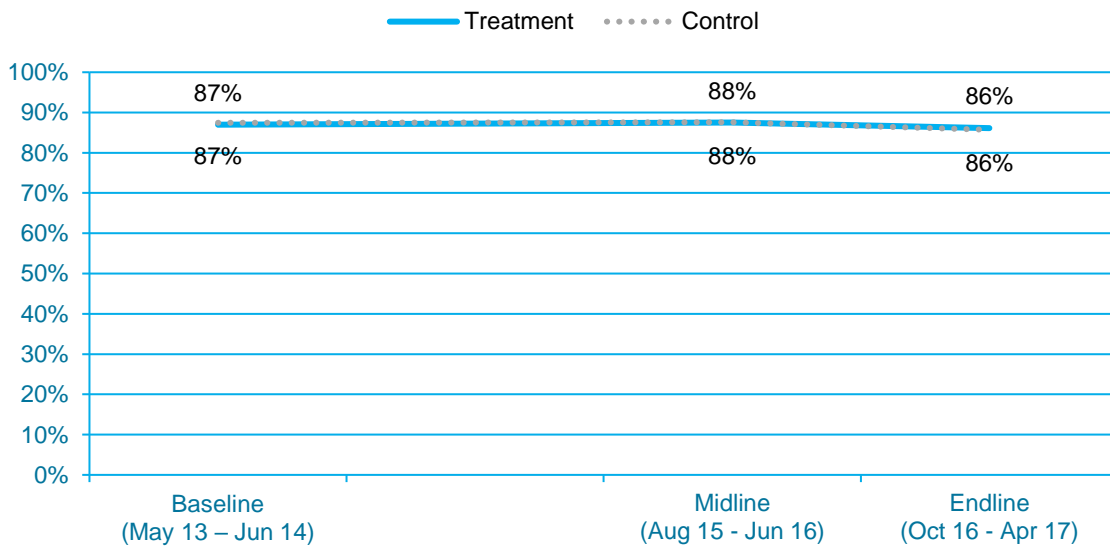
Across SCW, neither attendance measure showed a significant change from baseline, or from midline, compared to control areas. As shown in Figure 10, household-survey based attendance rates remained stable since baseline at around 86%. The rate is the same overall when calculated from the school visit survey, whether from the same set of categorical questions asked to the girl's teacher, or from the proportion of days attended when school was open. This suggests that **attendance levels remained approximately constant since baseline**.

At the project level, only AKF (Afghanistan) shows a positive and significant increase in attendance from the EM household-based measure (+7 pp). Most of the increase happened between midline and endline, but it could be mostly due to a catch-up effect of treatment areas who had substantially lower rates at baseline (83%) compared to control areas (89%).

3.3 What impact has the SCW had on marginalised girls' learning?

!

Difference-in-difference indicators from the EM data are followed by an asterisk if significant at the 10% level (p-value below 0.1) and by two asterisks if significant at the 5% level (p-value below 0.05). Tables showing difference-in-difference results across all EM data analytical variables are shown in Annex H.

Statistical significance depends not only on the size of the effect measured, but on the sample size on which this effect is calculated. For two effects of equal sizes, one can be significant if based on larger sample sizes, while the other is not significant. This feature needs to be kept in mind while comparing the significance of effects of two subgroups of very different sizes in our sample, such as in-school girls with out-of-school girls, or girls enrolled in primary school with girls enrolled in secondary.

Box 4: Notes on the analysis of learning data

The EM tests girls' learning in the household using the same learning assessment across project areas: The Early Grades Reading Assessment (for literacy) and Early Grade Mathematics Assessment (for numeracy).

The **version of EGRA**[58] used by the EM is made up of four subtasks: letter sound identification, invented word reading, oral reading fluency (of a passage /story) and reading comprehension. The first three subtasks are timed and students have a fixed time of 60 seconds to complete them. When completed in less than one minute, the time remaining is recorded. When students have attempted and failed the first n items of the subtask, with n equal to 10 per cent of the number of items, the subtask is discontinued by the assessor (except in the case of the reading comprehension subtask which is not timed).

The reading comprehension subtask is given only to girls who scored more than zero at oral reading. From midline, it is followed by two additional "B" oral reading and reading comprehension subtasks, which are more difficult, and are only given to girls who correctly answered three or more questions out of the five reading comprehension questions.

The **version of EGMA**[59] used is made up of five subtasks: number identification (timed), quantitative comparison, missing number identification, addition (timed) and subtraction (timed) plus four additional subtasks given only to the best-performing students (written exercise of addition, subtraction, multiplication and division).

The EM aggregates learning data across the SCW using a single learning metric for literacy and a single learning metric for numeracy:

- For EGRA: the **oral reading fluency** score is used which corresponds to the **number of words of a story correctly read in a minute**. This number is inflated by the time remaining at the end of the passage to create a **word-per-minute score**. The EGRA has other subtasks (letter reading, invented word reading, reading comprehension) whose scores are considered as well. But **the wpm score is the central benchmark used to track girls' progress in literacy** across the GEC.

- For EGMA: the **percentage of correct items across all subtasks** is used (EGMA/100). As with the wpm score, the time remaining at the end of each subtask is used to inflate the scores of students who reached the end of the subtask in less than one minute.

The use of time remaining allows us to remove ceiling effects from our literacy and numeracy score distributions. However, it implies that the best-performing students are differentiated mostly on the basis of their rapidity to finish a subtask.

The EGRA and EGMA are "early grade" assessments and may not be adapted to older students, especially those enrolled in secondary schools (who represent a large proportion of Camfed's sampled girls). It is therefore possible that the additional skills learned by students in higher grades may not be accurately captured by our measure of EGRA and EGMA.

---

[58] The EGRA toolkit can be found here: https://globalreadingnetwork.net/resources/early-grade-reading-assessment-egra-toolkit-second-edition
[59] The EGMA toolkit can be found here: https://globalreadingnetwork.net/resources/egma-toolkit

Girls are assessed in their own language of instruction, hence in different languages which have their own specificities and natural learning trajectories. All other things being equal, it may be more difficult, or take more years, for a child to read fluently in one language when compared to another. This may reduce the observable effect of projects aimed at improving learning in this language. Moreover, in some GEC project areas, a large number of girls do not speak the language of instruction as a native language. These restrictions need to be kept in mind when interpreting the findings described in the next section.

To gain a sense of how girls in SCW project areas fare in comparison with the reading fluency levels generally expected at a given age, **we refer to international benchmarks for oral reading fluency** published by Abadzi (2001)[60] for use by the World Bank. It is generally assumed that students must be able to read a minimum of 45-60 words per minute to be considered fluent readers and research suggests that this standard can possibly be applied worldwide[61].

3.3.1 What impact has the SCW had on literacy?

Impact across SCW

Average oral reading scores have increased across SCW target communities, although only marginally between midline and endline. At endline, girls scored 39 wpm on average, compared to 38 wpm at midline and 23 wpm at baseline (excluding Camfed areas). However, our analysis of EM data shows **no significant difference between the increases observed in treatment and in control areas**, as shown on Figure 11.

Figure 11: EGRA mean word-per-minute score across SCW project areas (excluding Camfed)
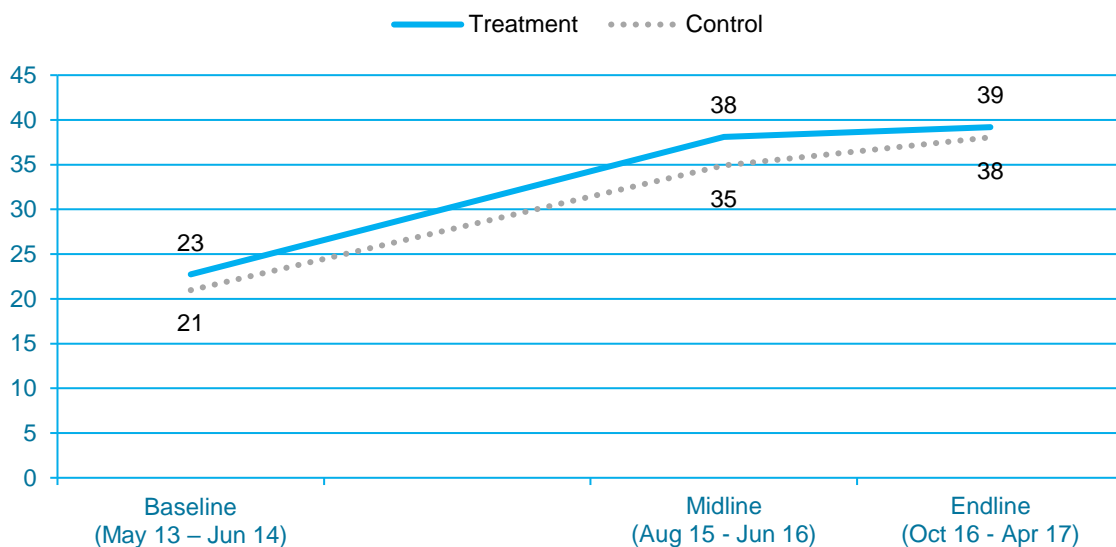
Impact by project area

In some project areas, literacy scores slightly decreased between midline and endline, while in others the pace of progress has kept about constant. Observed decreases are largely due to the small timespan between the midline and endline waves (less than twelve months) as well as due to measurement errors combined with the intrinsic variability of EGRA scores (the same student could score differently if she takes the test two successive times).

CfBT (Kenya) is the only project area which demonstrates a positive and significant increase in literacy between baseline and endline compared to control areas (+13 wpm*). Camfed demonstrates a similar significant increase (+15 wpm**) but their "control" areas are made of older girls, which tends to show that younger girls (aged 5-15 at baseline) have improved quicker than their older peers (aged 13-17 at baseline). It is worth noting that

---

[60] See Abadzi, H. (2011), *Reading Fluency Measurements in EFA FTI Partner Countries: Outcomes and Improvement Prospects*, GPE Working Paper Series on Learning, No. 1, Education for All Fast Track Initiative Secretariat, World Bank, Washington DC.
[61] Benchmark scores correspond to the scores achieved by the 18th percentile in US schools across different grades.

disparities in learning skills remain large across SCW treatment areas, with some girls scoring more than 70 wpm on average in some contexts, compared to less than 20 wpm in others (see Figure 12).

In Save the Children's (Ethiopia) target areas, we measure a negative midline to endline DiD (-2 wpm**) and a negative baseline to endline DiD (-3 wpm**).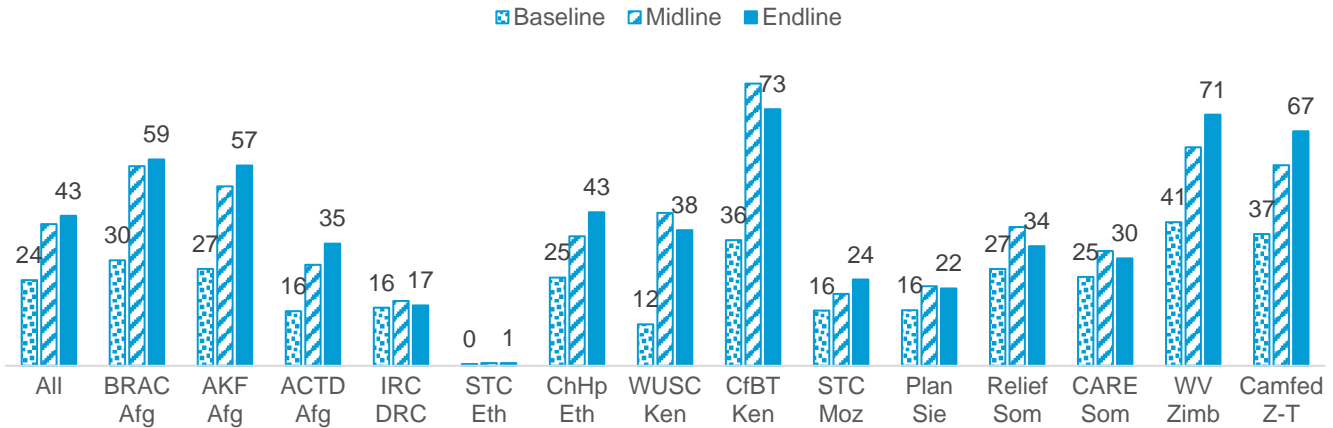 Indeed, girls' scores remained constant since baseline in treatment areas at 1 wpm on average while slightly increasing in control areas.

Figure 12: Mean oral reading score by SCW project treatment area[62]

Impact by subgroup

Older girls have achieved greater literacy gains than younger girls in absolute terms. In treatment areas, girls aged 9-11 and girls aged 12-13 improved by 3 wpm on average, while the oral reading score of girls aged 14-15 improved by an average of 6 wpm. However, we observe similar increases in control areas. Among school levels, upper primary girls have increased their scores significantly more than in control areas (+15 wpm**). This finding may be related to the fact that the girls most often exposed to education-related activities in SCW treatment areas are those enrolled in upper primary school, and consistent with the fact that it is the most targeted group across SCW (see Section 3.1). It is however worth noting that lower primary girls have been targeted by the same number of SCW projects, but do not show larger improvements than in control areas.

Out-of-school girls have improved by about 20 wpm, from 6 wpm to 27 wpm in treatment areas, **compared to 14 wpm for enrolled girls** who improved from 28 wpm at baseline to 42 wpm at endline, and 11 wpm for enrolled girls with low attendance. Girls who have never been enrolled also increased their scores by 20 wpm on average, although starting at a lower level (2 wpm at baseline) than girls who dropped out. Girls who were enrolled but had poor attendance increased their scores by only 11 wpm, from about 22 wpm to 34 wpm. This confirms the midline finding that **out-of-schools achieved comparatively greater learning gains than enrolled girls**, partly because they were lagging further behind than in-school girls from the same age so **their learning gains only brought them to the average starting point of the enrolled girls**. Evidence from projects at midline also suggested that in some contexts, it is easier to recreate the learning environment from scratch than to try improving learning conditions at school. All these increases have been comparable in control areas.

We observe **no significant effect among socioeconomic subgroups compared to control areas**, which suggests that although some subgroups may have been successfully reached, this did not translate into a larger impact on learning. It is worth noting though that all subgroups, except girls whose schooling is difficult to afford, increased their scores by 17 wpm to 19 wpm on average between baseline and endline, which is higher than the 16 wpm average increase across our SCW sample. All subgroups, except girls whose schooling is difficult to afford[63], started from a lower literacy level than the average girl at baseline. This confirms **that it is easier to**

---

[62] Values are expressed in words-per-minute. For the sake of clarity, midline value labels have been removed from the bar chart.
[63] This is because a large proportion of girls whose schooling is difficult to afford come from World Vision and Camfed areas in Zimbabwe where baseline literacy scores were relatively higher than elsewhere (see Table 8).

improve the score of girls who start from a lower basis. However, this finding is only true **for girls of the same age** as we showed above that older girls had greater learning gains than younger girls.

Impact by subtask

Table 11 shows the baseline to endline difference-in-difference coefficients of the EGRA subtasks, across the SCW project areas from the EM data. The first and easiest EGRA subtask, invented word, is the only subtask to show a positive and significant DID overall (+1.5 word*). The letter sound and oral reading subtasks show positive DID as well (+2.2 letters and +1.1 wpm respectively), but these are not significant.

As mentioned in Box 4, the reading comprehension subtask and the "B" oral reading and reading comprehension subtasks were only given to the best-performing girls. For the sake of comparability over time, girls who did not take these subtasks were assigned the score of zero. The two B subtasks, supposedly the hardest, exhibit negative difference-in-difference indicators, although only this B reading comprehension is significant (-0.1 question**).

Table 11: DID indicators by SCW project area for each EGRA subtask

Baseline to midline DID EM data	All excl. Camf.	BRAC Afg	AKF Afg	ACTD Afg	IRC DRC	STC Eth	ChHp Eth	WUSC Ken	CfBT Ken	STC Moz	Plan Sie	Relief Som	CARE Som	WV Zim	Camf. Z-T [2]
Invented Word	1.5*	0.5	4.2**	1.1	6.2*	-1.5	1.6	6.2	4.5	-1.8	-0.3	6.8	-0.6	-2.6	7.3
Letter Sound	2.2	0.7	6.9**	1.7	0.8	-2.7	2.9	-0.3	7.2**	1.1	-2.1	0	4	1	0.7
Oral Reading	1.1	0.9	4.4	3.6	2.8	-2.8**	1.1	8.8	13.2*	-2.4	-3.4	5.4	-2.6	-3.5	15
Reading Comprehension	0.1	-0.1	0.3	0.3	0	-0.1	0.1	0.3	0.4	-0.1	-0.2	-0.1	-0.2	0	0.8
B - Oral Reading [1]	-1.8	-1.3	-6.9**	-2.3	-0.2	-0.7	0.8	1.2	2.4	1.9	-1.9	-3.1	-18.2*	0.7	17.9
B - Reading Comprehension [1]	-0.1**	-0.1	-0.3**	0	0	0	-0.1	-0.1	0.1	0.1	0	-0.1	-0.8**	0	0.2

Note: (1) B subtasks were introduced at midline, and given only to best-performing students. Other girls were assigned a zero score for this subtask. For these two subtasks, we show the difference-in-difference indicator between midline and endline.
(2) Camfed areas are included for indicative purposes but do not have a proper control group (see explanation at the beginning of Section 3.2).
Light green = score increased significantly at the 10% level. Light orange = score decreased significantly at the 10% level.
Darker green = score increased significantly at the 5% level. Darker orange = score decreased significantly at the 5% level.

Overall, **only the easiest subtasks show a positive effect** in treatment compared to control areas. Conversely, the most difficult subtasks tend to show negative difference-in-difference. A striking example of this is AKF (Afghanistan) project areas where girls significantly increased their scores in invented word and letter sound, had no effect on oral reading and reading comprehension, but did less well than control areas when it came to the more difficult B subtasks. This is not due to girls in AKF (Afghanistan areas) having started lower than in other contexts: at baseline they scored 27 wpm on average in AKF treatment areas, compared to 23 wpm for the average girl across SCW.

Endline wpm compared to international benchmarks

Figure 13 shows the oral reading scores achieved by girls across the SCW (treatment and control areas) as blue bars, depending on how many years of school they attended (horizontal axis). The orange dots show the Abadzi international benchmarks for developing countries (see Box 4).

Figure 13: Endline wpm scores by number of years in school and corresponding benchmarks
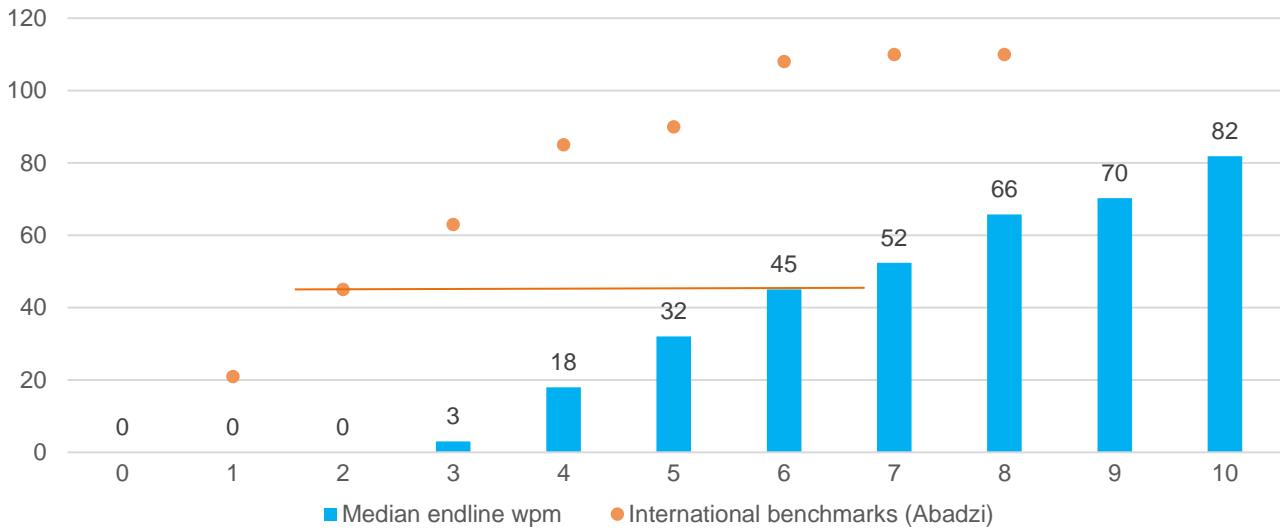
The horizontal line shows the "basic comprehension" threshold of 45 wpm that is supposed to be attained by students during their second year in primary school. It corresponds to the minimum skill level that allows students to transition from learning to read, to reading to learn. As shown on the figure, at endline this level is only reached by girls who attended six years of school, which for most girls corresponds to the end of primary school. This suggests that **girls in SCW areas are still considerably lagging behind international benchmarks**, at the risk of affecting not only their literacy and numeracy skills but the pace at which they are learning all other subjects.

It is worth noting that the progress made by GEC projects was measured as improvements from the baseline learning levels of their target girls. Projects' achievements were not assessed against benchmarks of absolute learning levels. At baseline, it was found that learning levels were lower than what might have been originally expected. This largely explains the prevailing low levels of learning across SCW areas.

3.3.2 What impact has the SCW had on numeracy?

Numeracy scores show a similar trend as literacy across the SCW with a continuous increase from baseline to endline from 40% at baseline to 60% at midline and 65% at endline. The trajectory is **almost identical in treatment and control areas**.

Figure 14: EGMA % correct across SCW project areas (excl. Camfed)
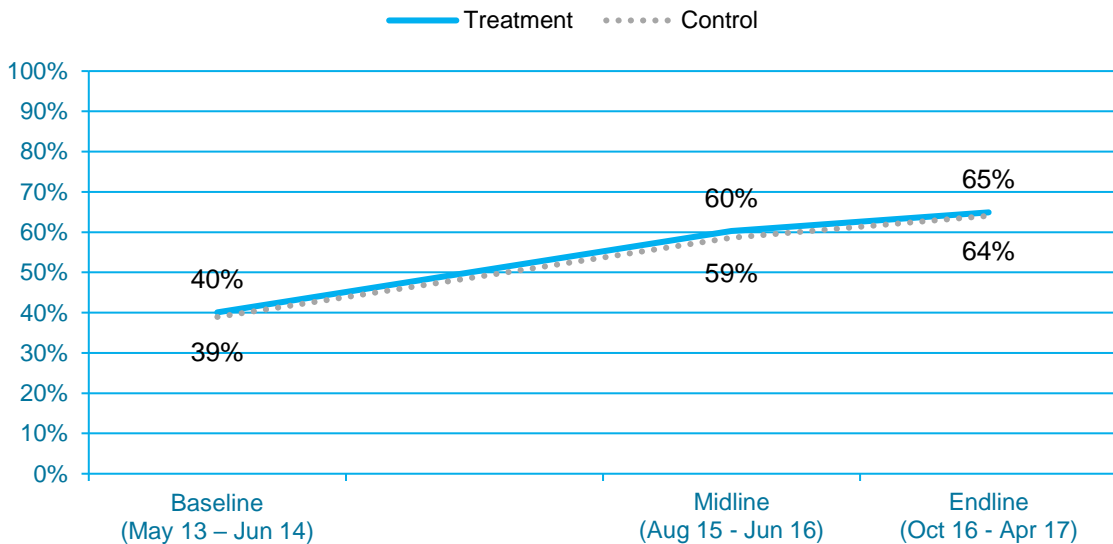
AKF (Afghanistan) is the only SCW project area showing a positive and significant DiD between baseline and endline (+9 pp*). Most of this increase has been realised between baseline and midline (+7 pp*). Conversely, the DiD is negative in Save the Children (Ethiopia) areas (-8 pp**), as it was for literacy. **Score patterns are equivalent for literacy and numeracy:** project areas where girls score the highest in EGRA are the same where they score the highest in EGMA, and vice versa. At endline across SCW, the linear correlation between the percentage of questions correct in EGRA and EGMA is equal to 80%.

As for the oral reading score, the **EGMA score shows a positive and significant improvement since baseline for the age category of 9-11-year-old girls only, compared to control areas**. This confirms that younger girls have been most effectively reached by SCW projects. However, looking at improvements across time in treatment areas, the trend is the same as for literacy, with older girls improving more than younger girls: in treatment areas, girls aged 14-15 increased their numeracy score by 11 pp on average between baseline and endline, girls aged 12-13 by 8 pp, girls aged 9-11 by 2 pp and girls aged 6-8 by less than 1 pp. These increases are similar in control areas, which seems to confirm the literacy finding that generally, older girls increase their scores more easily than younger girls (irrespective of whether they belong to treatment or control communities).

However, none of the Lower Primary, Upper Primary or Secondary phase categories show a significantly better improvement than in control areas. As for oral reading, enrolled girls have improved less than out-of-school girls over time: from 48% to 69% (+21 pp) for in-school girls in treatment areas compared to 12% to 46% (+34 pp) for out-of-school. This again confirms the finding from the literacy assessment that **learning gains over time have been greater for out-of-school girls than for in-school girls**, although gains are not significant for both groups.

For socioeconomic subgroups, findings are similar for oral reading fluency: EGMA scores did not improve more in treatment than in control areas, but in treatment areas, improvements over time were on average slightly higher for girls belonging to subgroups than for the average SCW girl. Again, this confirms the literacy finding that, **in absolute terms, it is easier to improve the learning of low-performing girls than that of other girls of the same age.**

Table 12 shows the difference-in-difference indicators of EGMA subtasks, starting with the first and easiest syntax at the top, up to the more difficult written exercise tasks, which have been given only to girls who correctly solved five or more addition or subtraction level 1 items. As opposed to EGRA, we do not observe larger improvements in easier subtasks compared to more difficult ones. Save the Children (Ethiopia) even presents a reverse pattern, with significant negative DiD for the three first subtasks. The absence of a significant decrease for higher subtasks in this project area is largely due to the facts that most girls scored zero at baseline, hence could not score lower.

The intermediary addition and subtraction subtasks are the tasks where girls made the greatest progress across the SCW. In Relief International's (Somalia) target areas, the two scores increased significantly by 7.2 items and 3.7 items respectively. Overall, addition level 1 is the only subtask showing a positive and significant DiD (+0.8 item*). As for EGRA, Camfed shows large and significant improvements across all subtasks, demonstrating that younger girls (Walk A) have improved their skills more than older girls (Walks B + C).

Table 12: DiD indicators by SCW project area for each EGMA subtask

Baseline to midline DID EM data	All excl. Camf.	BRAC Afg	AKF Afg	ACTD Afg	IRC DRC	STC Eth	ChHp Eth	WUSC Ken	CfBT Ken	STC Moz	Plan Sie	Relief Som	CARE Som	WV Zim	Camf. Z-T [1]
Number identification	0.5	0.3	1.5	0.1	1.3	-5.6**	3.3	0.8	2.3	-0.1	-0.4	0.5	4.6	-1.9	11.8
Quantity Comparison	0.1	-0.2	0.5	0.3	-0.7	-0.6*	0.9**	-0.1	1	0.5	-0.3	-0.6	-0.3	-0.6	2.7
Missing Number	0.1	0.5	0.5	0.6	-0.6	-0.5**	0.4*	0.3	0.5	-0.2	-0.7	0	-0.6	-0.1	2.4
Addition Level 1	0.8*	0.7	2	0.5	2.1	-0.5	1	3*	2.2*	-0.9	-0.9	7.2*	-1.8	-1.4	2.9
Subtraction Level 1	0.4	0.7	2.3**	0.7	1.6	-0.4	0.1	2.2	0.9	-1.8	0.9	3.7*	-4.6	-0.7	2.9
Written Exercise - Addition	0	0	0.1	0.3	-0.2	-0.1	0.2	0.1	0	-0.3	-0.2	0.3	-0.6*	-0.2	0.9
Written Exercise - Subtraction	0.1	0.1	0.8**	0.3	0	-0.1	0	0.3	0.1	-0.5**	-0.2	0.4	-0.6*	-0.2	1.0
Written Exercise - Multiplication	0	0.1	0.3	0	0.1	-0.1	0.1	0.3	0.2	-0.2	0.1	0.2	-0.3	0	0.8
Written Exercise - Division	0.1	0.2	0.3	0	0	0	-0.1	0.4	0.4	-0.2	0.2	0.2	-0.4	0.1	0.7
% Correct across all subtasks	2%	2%	9%*	3%	4%	-8%**	6%	8%	8%	-4%	-1%	12%	-5%	-5%	27%

Figure 15 shows the endline median EGRA oral reading score (left vertical axis) and EGMA/100 score (right vertical axis) across the SCW, by categories of school attendance. Here, attendance is measured in school as the ratio of the number of days a girl was present over the number of days the school was open. As mentioned in Section 3.2.2, this measure is only available in areas where a school visit survey has been conducted by the EM.

Figure 15: Endline median learning scores across SCW project areas by attendance in school
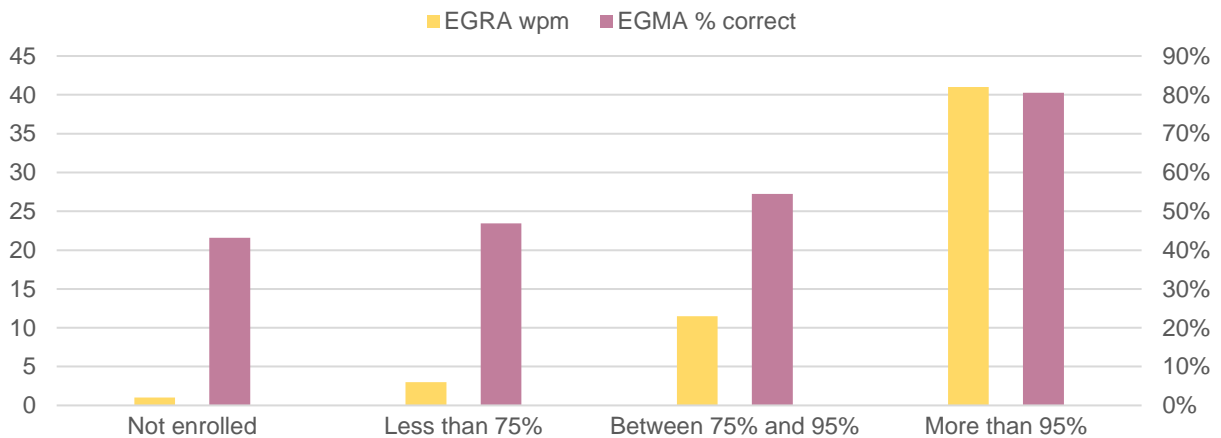
Literacy scores show a much sharper difference across attendance levels than numeracy scores. It is striking to note that girls who attend more than 95% of days can read forty times more words-per-minute than non-enrolled girls, and twenty times more than enrolled girls with poor attendance, but score no more than twice as much in EGMA than girls in these two categories. Some of this variation may be due to the scoring of the EGRA / EGMA tests: girls with very low formal numeracy skills are able to correctly answer a range of EGMA items because even girls who have never or only rarely attended school may have acquired basic numeracy skills from their everyday life, while many girls among SCW target communities still cannot read a single word. However, it also confirms the midline finding that **girls tend to progress less in numeracy than literacy when they are in school**.

3.3.3 Difference-in-difference on deciles of the learning scores' distribution

Box 5: Deciles of a distribution

Deciles split the score distribution into ten equal parts. For example, the lowest 10% of girls range from the smallest score to the first decile (D1). The next 10% has scores between D1 and D2. Scores of the highest ten percent range from the ninth decile (D9) to the highest score. The intermediary 5th decile (D5) is also called the median, which is the score value at which the distribution is split within two equal parts.

Deciles should be interpreted in the following way: if the 7th decile of the oral reading score distribution is 60 words per minute, then a girl scoring 60 words scores better than 70% of the test-takers. If the median of the oral reading score is 20 words per minute, it means that half the girls score less than 20 words per minute, and half the girls score more than 20 words per minute.

The advantage of deciles is that they provide information not only about the 'average' student, but about different parts of the learning score distribution, from the lowest to the highest performing students. Besides, as opposed to the mean, they are not sensitive to large and extreme values. The oral reading score ranges from 0 wpm to 200 wpm (a few girls scored higher than 200 wpm after accounting for time remaining, but their scores were capped at 200). However, there are a lot more girls scoring around 0 wpm than around 200 wpm. The mean may therefore give a wrong sense of the scale of the effect observed. For instance, the mean difference is the same when 100 girls progress from 0 to 10 wpm as when 10 girls progress from 100 to 200 wpm.

Figure 16 shows the baseline to endline difference-in-difference indicators for the EGRA oral reading score and the EGMA aggregate score across the nine deciles of their respective distributions, across all SCW project areas (excluding Camfed). Both scores are shown on the same scale: the EGRA score is expressed in words-per-minute, and the EGMA /100 score. For EGMA, a DiD of +1.6 therefore means that the EGMA score has increased by 1.6 pp more in the treatment group than in the control group. Asterisks show statistical significance: one asterisk shows significance at the 10% level, two asterisks show significance at the 5% level.

Figure 16: Baseline-endline difference-in-difference of EGMA/EGMA scores by decile across SCW project areas (cohort excluding Camfed)
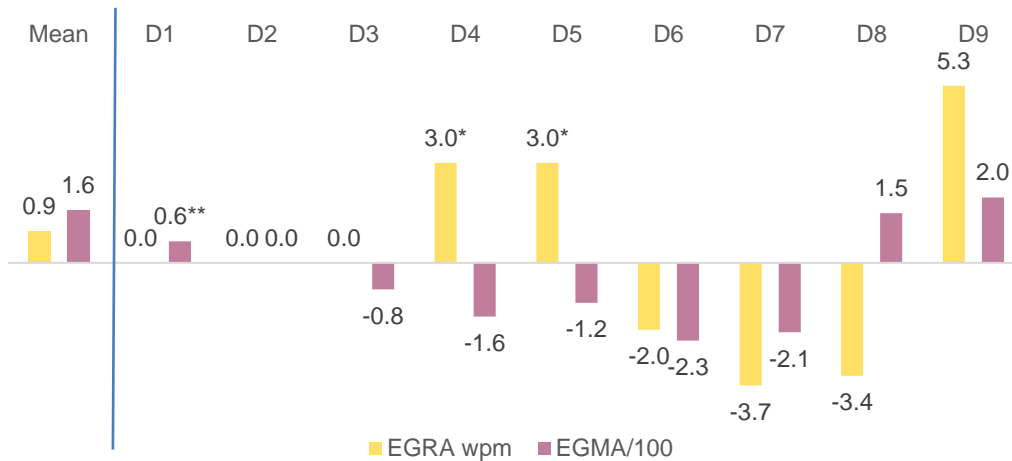
The EGRA score is significantly higher in treatment areas, by +3 wpm, for two deciles: the fourth decile (D4) and the fifth decile (D5), the median. This confirms the midline finding that the **SCW impacted the most on literacy scores around the middle of the distribution**, rather than on the lowest or highest performing girls in the cohort.

The EGMA score is significantly higher only for the lowest decile. It should be noted that despite the low absolute size of the effect (+0.6 pp), it is significant at the 5% level. This reflects the fact that students from the lowest decile have very low scores: hence, an increase of +0.6 pp is relatively large for them. This again confirms the midline finding that the **overall effect of SCW on numeracy has tended to focus on the lowest-performing girls**, rather than on the medium deciles (as it did for literacy).

All other deciles show non-statistically significant DiD at either the 5% or the 10% level.

Although these findings seem to somewhat contradict the EGRA and EGMA subtask analysis presented earlier, they offer some interesting nuances:

- For **EGRA**, we noticed that the easiest subtasks show the largest improvements compared to control areas. Here we demonstrate that the oral reading score, which does not show a significant improvement overall, does show the largest improvements around the middle of the distribution. This in particular suggests that girls from the first half of the distribution (lowest performers) did not increase their scores enough. Taking a closer look, D3 is still equal to 0 wpm at endline (as well as D1 and D2) in treatment areas. This means that **at least 30% of girls across the SCW target communities still cannot read a single word.** The score of these girls did not improve between baseline and endline, which explains why we see the relatively largest increase for D4 and D5 which, excluding zero scores, are the "second-lowest deciles" of the oral reading score distribution.

- For **EGMA**, the largest (relative to the baseline score) and only significant improvement is observed for the lowest decile of the distribution. We presented earlier that intermediary subtasks (addition and subtraction) show the largest improvements compared to control areas. This could suggest that the lowest performing girls have improved more in the addition and subtraction subtasks, which is an encouraging finding.

3.3.4   Outcomes by exposure

We further explore the change in learning outcomes based on whether or not girls have been exposed to specific GEC interventions. Table 13 shows the change in EGRA and EGMA scores between baseline and endline, comparing girls who have been exposed to intervention (according to their caregiver, teacher or head teacher) with girls who have not.

Table 13: Baseline to endline changes in learning score, by exposure

EM data		Receiv. school books	Received a scholarship or bursary		Attended special classes or study groups		Received special tutoring / help with schoolwork		Talked enrolling / staying in school	NGO / relig org° supplied funds	Activ. to improve learning at school	Ongoing measures to support girls	Teachers were trained in: General teaching methods	Gender sensitive pedago.	Special needs educ.	Mentor. or coaching
		HHS	HHS	SVS	HHS	SVS	HHS	SVS	HHS	SVS	SVS	SVS	SVS	SVS	SVS	SVS
EGRA wpm - oral reading score																
Treatment	No	22	22	20	22	21	20	19	22	22	14	13	12	18	18	14
	Yes	25	23	22	24	22	31	27	24	19	23	24	23	23	29	27
	Diff.	3	1**	1	2	1	11**	7	2	-2	9**	11**	11**	5	11*	13**
Control	No	19	18	20	20	22	18	23	19	21	15	15	9	18	18	16
	Yes	21	24	25	18	21	24	21	19	15	21	21	22	20	23	22
	Diff.	2	6**	5	-2	-1	5**	-1	0	-6	6**	5**	13**	1	5*	7**
EGMA/100 - % questions correct																
Treatment	No	28%	28%	31%	28%	31%	27%	29%	29%	33%	25%	27%	26%	32%	30%	29%
	Yes	39%	34%	39%	35%	35%	38%	38%	31%	33%	35%	35%	35%	33%	39%	36%
	Diff.	11%	6%	7%*	7%	4%**	11%**	8%	3%	0%	9%**	8%**	9%**	1%	9%	7%**
Control	No	24%	25%	31%	24%	30%	24%	32%	25%	30%	22%	24%	22%	31%	30%	28%
	Yes	27%	19%	40%	27%	35%	28%	32%	24%	31%	33%	33%	33%	30%	31%	33%
	Diff.	2%	-6%	9%*	3%	6%**	4%**	0%	-1%	1%	11%**	9%**	11%**	-1%	1%	5%**

Note: "Yes" rows correspond to the average change across girls who have been exposed to activities. "No" corresponds to those who haven't. "Diff" rows show the difference between "Yes" rows and "No" rows. They are coloured in dark green when the p-value of the t-test Yes < No is lower than 5%, and coloured in light green when it is lower than 10%. Calculations made on girls who have been tracked from baseline to endline. HHS refers to EM Household-Survey data. SVS refers to School-Visit-Survey data, in the countries where school visits took place.

We observe **little difference overall between treatment and control areas**. This is very likely to be due to GEC and non-GEC activities taking place in control areas, as well as some projects not working any more in schools they identified before baseline as belonging to their treatment group (see Section 2 for more details on the limitations of the EM sample). It is therefore impossible to disentangle the effects of specific GEC activities from the effects of similar activities taking place in the same areas. However, the above table provides interesting insights about the types of activities that may have had a larger impact on learning, most of which have a positive effect on both numeracy and literacy.

Girls who received a scholarship or bursary between baseline and endline improved their learning more than those who did not. But the difference in improvements is relatively small (+1 wpm in treatment areas, +6 wpm in control areas). Note that the differences were calculated including Camfed areas, where control areas contain girls who did receive a bursary from the project. When excluding them, the effect across control areas decreases to +2 wpm and is not significant.

When activities to improve learning and to support girls took place in girls' schools their learning improved significantly more since baseline. The literacy scores show a larger difference in treatment areas for "Activities to improve learning at school" and "Ongoing measures to support girls" (+9 wpm and + 11 wpm respectively) than in control areas (+6 wpm and + 5 wpm), which tends to suggest that SCW projects had a special focus on literacy, or have been relatively more effective in improving literacy. The numeracy scores show significant differences too, but these are comparable across treatment (+9 pp and +8 pp) and control areas (+11 pp and +9 pp).

Looking at specific activities, **attendance of classes or study groups has had a small but significant effect on girls' EGMA scores** (+4 pp in treatment areas, +6 pp in control areas). This seems to confirm the findings from some SCW projects that **numeracy is more difficult to teach and learn than literacy and requires smaller groups and a specific focus to help girls make substantial progress.**

Similarly, **girls who received special tutoring or help with their schoolwork have improved significantly more in both literacy and numeracy than those who have not**. Besides, differences are much larger in treatment areas (EGRA: +11 wpm, EGMA: +11 pp) than in control areas (EGRA: +5 wpm, EGMA: +4 pp), suggesting that such SCW activities have been particularly effective in improving girls' learning.

In terms of teachers' training, **two types of training seem to have been particularly effective at improving girls' learning in both literacy and numeracy: general teaching methods and mentoring or coaching**. Conversely, training teachers in gender sensitive pedagogy and in special needs education does not (in itself) seem to have positively impacted on girls' learning between baseline and endline. These trainings might still have had a significant effect on specific subgroups of girls, but our sample sizes are not large enough to capture it.

This analysis suggests **that interventions that aim at improving learning directly, such as tutoring, learning support or teacher training, have had a larger positive effect on girls' learning than indirect "activities"** such as the provision of school materials, scholarships, talks about enrolling or staying in school, or trainings in gender sensitive pedagogy. Some of these interventions may help girls go to school and stay in school as well as create better conditions for learning, but they have not proved sufficient to significantly increase girls' learning between baseline and endline. The effectiveness of different types of intervention will be discussed in greater details in Section 3.4.

3.3.5 SCW projects' learning achievements

⚠️ As in the previous sub-section, the analysis and interpretation described here are based on SCW projects' own evaluation data, as opposed to findings presented in sub-sections 3.3.1 to 3.3.3, which were based on EM data analysis.

Table 14 below shows the SCW projects' achievements in literacy using projects' own data as shown in the outcome spreadsheets and validated by the FM.

Table 14: SCW projects' achievements in literacy based on projects' own data reanalysed by the FM

SCW projects' achievement	BRAC-C	BRAC-G	AKF	Acted	IRC	STC	ChHpe	WUSC	CfBT	STC	Plan 1	RI	CARE	WV	Camfd	Camfd
	Afg	Afg	Afg	Afg	DRC	Eth	Eth	Ken	Ken	Moz	Sie	Som	Som	Zim	Tan	Zim
Midline target in SD	0.3	0.3	0.3		0.3	0.3	0.3	0.2	0.25	0.25		0.2	0.2	0.2	0.3	0.3
Midline data quality	Partly conclusive	Partly conclusive	Partly conclusive		Partly conclusive	Partly conclusive	Conclusive	Conclusive	Conclusive	Partly conclusive		Partly conclusive	Partly conclusive	Partly conclusive	Partly conclusive	Partly conclusive
Endline target in SD	0.2	0.2	0.3	0.3*	0.2	0.2	0.2	0.2	0.2	0.25	0.2*	0.3	0.3	0.2	0.2	0.2
Endline data quality	Partly conclusive	Partly conclusive	Conclusive	Partly conclusive	Conclusive	Inconclusive	Conclusive	Partly conclusive	Conclusive	Partly conclusive	Inconclusive	Partly conclusive	Partly conclusive	Conclusive	Partly conclusive	Partly conclusive
Control group	No	No	No	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Literacy test and unit	EGRA wpm	EGRA wpm	EGRA wpm	ASER	EGRA /100	EGRA /100	EGRA wpm	EGRA wpm	UWEZO	EGRA /100	ASER /10	Adapted UWEZO	Adapted UWEZO	EGRA wpm	National	National
Baseline (Treatment)	44	77	16	1.6	11	10	40	66	2.8	28	1.9	6.6	38	33	N/A	N/A
Baseline (Cont./Bench.)	44	88	24	N/A	8	18	40	72	2.7	31	1.8	8.8	61	35	N/A	N/A
Midline (Treatment)	66	95	44	N/A	23	18	52	103	3.9	76	N/A	7.3	61	70	40	27
Midline (Cont./Bench.)	45	89	22	N/A	16	34	53	72	3.6	79	N/A	7.4	65	73	38	31
Endline (Treatment)	96	120	67	3.2	38	24	53	100	4.6	61	4.8	8.2	80	77	53	36
Endline (Cont./Bench.)	69	99	45	1.8	24	23	54	99	3.9	63	5.0	8.1	72	76	40	35
BL-ML Impact in % (adjusted)	218%	8%	240%	N/A	162%	-153%	12%	299%	99%	23%	N/A	-15%	178%	24%	215%	96%
Impact in SD (adjusted)	0.65	0.02	0.72	N/A	0.49	-0.46	0.04	0.60	0.26	0.06	N/A	-0.03	0.36	0.05	0.65	0.29
ML-EL Impact in % (adjusted)	539%	333%	296%	453%	115%	461%	10%	8%	144%	24%	-62%	5%	134%	11%	465%	146%
Impact in SD (adjusted)	1.08	0.67	0.89	1.36	0.23	0.92	0.02	0.02	0.29	0.06	-0.12	0.02	0.40	0.02	0.93	0.29
BL-EL Impact in % (adjusted)	346%	138%	268%	453%	143%	93%	11%	154%	119%	24%	-62%	-3%	152%	18%	315%	116%
Impact in SD (adjusted)	1.73	0.69	1.61	1.36	0.72	0.46	0.06	0.61	0.53	0.12	-0.12	-0.01	0.76	0.07	1.58	0.58

Description of the table

The rows "Midline target in SD" and "Endline target in SD" show the projects' targets in standard deviations, over and above the control group (or benchmarks for projects which did not use control groups, hence the scores designated "Cont./Bench"), from baseline to midline and midline to endline respectively.

Benchmarks were set by the FM in agreement with the project, and were usually defined as the average baseline (for midline impact) or midline (for endline impact) scores of girls enrolled in one or two grades higher than the cohort. For instance, treatment girls in grade 6 at midline would have their endline scores compared with those achieved by girls who were in grade 7 at midline.

When the endline target is marked with an asterisk (i.e. ACTED and Plan), it denotes a baseline to endline target (as opposed to a midline to endline target) because these projects did not submit any learning data at midline. Note there are two columns for BRAC (Afghanistan), because community-based schools and government schools are assessed separately, and two for Camfed's areas covering Tanzania and Zimbabwe separately.

A target equal to 0.2 implies that the difference-in-difference (DID) coefficient must be equal to 0.2 times the standard deviation for the project to show a 100 per cent impact on its target. In other words, the "impact in %" is equal to the ratio of the DID coefficient over the target. When the project does not have control groups, benchmarks are set based on the scores obtained by higher grade students enrolled one or two years above.

"Midline data quality" and "Endline data quality" show the FM's rating of the quality of the learning data submitted by the project. These ratings are shared between the literacy and numeracy scores and were reached by judging the quality of data against comparability, compliance, contamination, completeness and clarity. In particular, "Partially Conclusive" data shows a balance of quantitative and qualitative evidence that gives some evidence of progress, but the FM cannot be confident in attributing the extent of project effect, and results are often statistically insignificant.

The type of learning assessment and unit used are shown at the top of the literacy and numeracy tables. The following rows show baseline, midline and endline scores for the treatment and control groups using projects' own learning data validated by the FM. They are weighted averages extracted from the outcome spreadsheets. These averages have been calculated across different grade levels or girls' populations (in-school versus out-of-school). The weighting used is proportional to sample size and should approximately reflect each project's beneficiary population.

The "impact in %" rows show the final impact figures for each project, corresponding to the percentage of target achieved. These figures were agreed with the projects and validated by the FM. For projects using control groups, this figure was obtained through a difference-in-difference regression analysis. The FM then adjusted this regression by refining the model or using control and clustering variables to account for the specificities of each project's evaluation design. The baseline to endline impact has been calculated by the EM as the sum of the baseline to midline and midline to endline impacts, over the sum of the two targets.

Colours reflect the size of the impact: red when negative or when data is inconclusive, yellow when positive but lower than 100%, light green between 100% and 150%, darker green when higher than 150%.

Figure 17 below shows aggregated baseline, midline and endline literacy scores for the treatment group as well as projects' literacy achievements. Figures in circles show the endline literacy impact in %. Circles are filled with the corresponding colour (endline impact). Their outline is coloured according to the size of the midline impact.

Figure 17: SCW projects' final impact and aggregated literacy scores for the treatment group
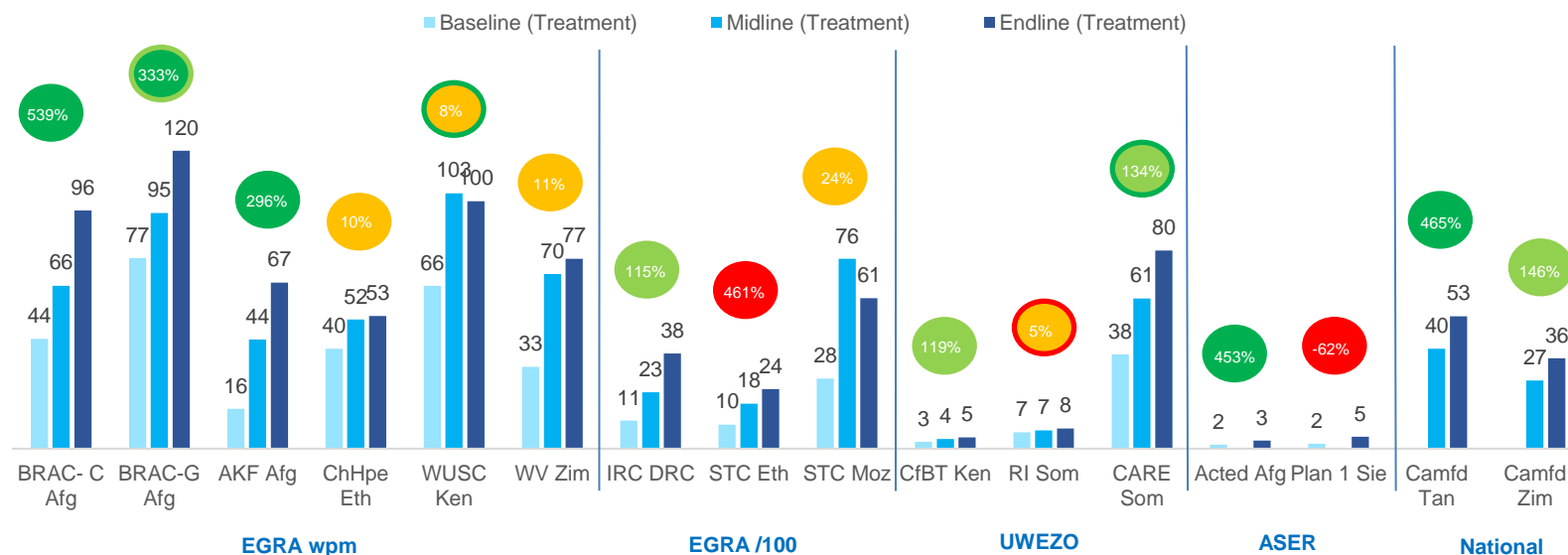
Interpretation

According to the FM's reanalysis of project data, half of the projects (eight out of 16) have met or exceeded their endline literacy target. Out of these eight projects, five did not have control groups and were assessed using benchmark scores. The projects without control groups that did not meet their targets are WUSC (Kenya) and RI (Somalia). Projects that did not meet their targets are those which show a decrease in the scores of their treatment group (STC (Mozambique) and WUSC (Kenya)) or a marginal increase since midline (ChildHope (Ethiopia), World Vision (Zimbabwe) and Relief International (Somalia)).

Looking across different types of learning metrics and tests used, the picture is *balanced*: four projects out of the nine projects who used EGRA wpm scores met their endline targets. Two projects out of the three using UWEZO met their endline targets, and one project out of the two who used ASER met their endline targets (note these two projects did not have midline learning data and were therefore assessed between baseline and endline). Camfed (Tanzania and Zimbabwe) used a national assessment and exceeded their target in both countries.

Endline achievements are usually correlated with midline achievements: the majority of projects who met their midline targets also met their endline targets, and projects who fell short of achieving their midline targets usually did not meet their endline targets. Two contradictory examples to this are BRAC-G (Afghanistan) and WUSC (Kenya). BRAC-G (Afghanistan) achieved just 8% of their midline target but 333% of their endline target. Conversely, WUSC (Kenya) achieved 299% of their midline target but only 8% of their endline target. These two projects used benchmarks, it is therefore possible that this inconsistency is linked to the way treatment scores were compared against benchmark scores.

Save the Children (Ethiopia) has achieved its target by 461% at endline compared to control areas after a very large decrease between baseline and midline (-153%), but these figures are considered inconclusive due to serious concerns around the quality of learning data. Similarly, Plan (Sierra Leone) baseline to endline achievement is negative but inconclusive due to uncertainties around the tracking of the treatment group.

Table 15 below shows the projects' achievements in numeracy using projects' own data as shown in the outcome spreadsheets and validated by the FM.

Table 15: SCW projects' achievements in numeracy based on projects' own data reanalysed by the FM

SCW projects' achievement	BRAC- C	BRAC-G	AKF	Acted	IRC	STC	ChHpe	WUSC	CfBT	STC	Plan 1	RI	CARE	WV	Camfd	Camfd
	Afg	Afg	Afg	Afg	DRC	Eth	Eth	Ken	Ken	Moz	Sie	Som	Som	Zim	Tan	Zim
Midline target in SD	0.3	0.3	0.3		0.3	0.3	0.3	0.2	0.25	0.25		0.2	0.2	0.2	0.3	0.3
Midline data quality	Partly conclusive	Partly conclusive	Partly conclusive		Partly conclusive	Partly conclusive	Conclusive	Conclusive	Conclusive	Partly conclusive		Partly conclusive	Partly conclusive	Partly conclusive	Partly conclusive	Partly conclusive
Endline target in SD	0.2	0.2	0.3	0.3*	0.2	0.2	0.2	0.2	0.2	0.25	0.2*	0.3	0.3	0.2	0.2	0.2
Endline data quality	Partly conclusive	Partly conclusive	Conclusive	Partly conclusive	Conclusive	Inconclusive	Conclusive	Partly conclusive	Conclusive	Partly conclusive	Inconclusive	Partly conclusive	Partly conclusive	Conclusive	Partly conclusive	Partly conclusive
Control group	No	No	No	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	No	Yes	Yes	Yes
Numeracy test and unit	EGMA / 100	EGMA / 100	EGMA / 100	ASER	EGMA / 100	EGMA / 100	EGMA / 100	EGMA / 100	UWEZO	EGMA / 100	ASER / 22	Adapted UWEZO	Adapted UWEZO	EGMA / 100	National	National
Baseline (Treatment)	50	61	26	2.6	47	17	50	79	N/A	32	9.7	6.0	36	50	N/A	N/A
Baseline (Cont./Bench.)	50	66	34	N/A	42	40	53	82	N/A	34	9.3	8.6	60	51	N/A	N/A
Midline (Treatment)	55	63	50	N/A	52	26	64	78	44	87	N/A	6.1	51	67	14	16
Midline (Cont./Bench.)	50	66	31	N/A	47	38	63	82	44	86	N/A	6.2	61	69	9	20
Endline (Treatment)	66	71	58	5.1	65	37	68	80	49	76	15.1	6.8	67	71	33	25
Endline (Cont./Bench.)	57	68	51	2.9	54	34	68	77	44	76	15.2	6.9	65	72	14	27
BL-ML Impact in % (adjusted)	87%	-167%	290%	N/A	69%	151%	28%	-116%	N/A	42%	N/A	-171%	-17%	68%	511%	88%
Impact in SD (adjusted)	0.26	-0.50	0.87	N/A	0.21	0.45	0.08	-0.23	N/A	0.11	N/A	-0.34	-0.03	0.14	1.53	0.26
ML-EL Impact in % (adjusted)	275%	126%	137%	420%	127%	290%	-10%	114%	78%	-3%	-38%	-21%	35%	-4%	501%	114%
Impact in SD (adjusted)	0.55	0.25	0.41	1.26	0.25	0.58	-0.02	0.23	0.16	-0.01	-0.08	-0.06	0.10	-0.01	1.00	0.23
BL-EL Impact in % (adjusted)	162%	-50%	214%	420%	92%	207%	13%	-1%	78%	20%	-38%	-81%	14%	32%	507%	99%
Impact in SD (adjusted)	0.81	-0.25	1.28	1.26	0.46	1.03	0.06	0.00	0.16	0.10	-0.08	-0.40	0.07	0.13	2.54	0.49

Figure 18 below shows aggregated baseline and midline numeracy scores for the treatment group as well as projects' final impact in percentage terms.

Figure 18: SCW projects' final impact and aggregated numeracy scores for the treatment group
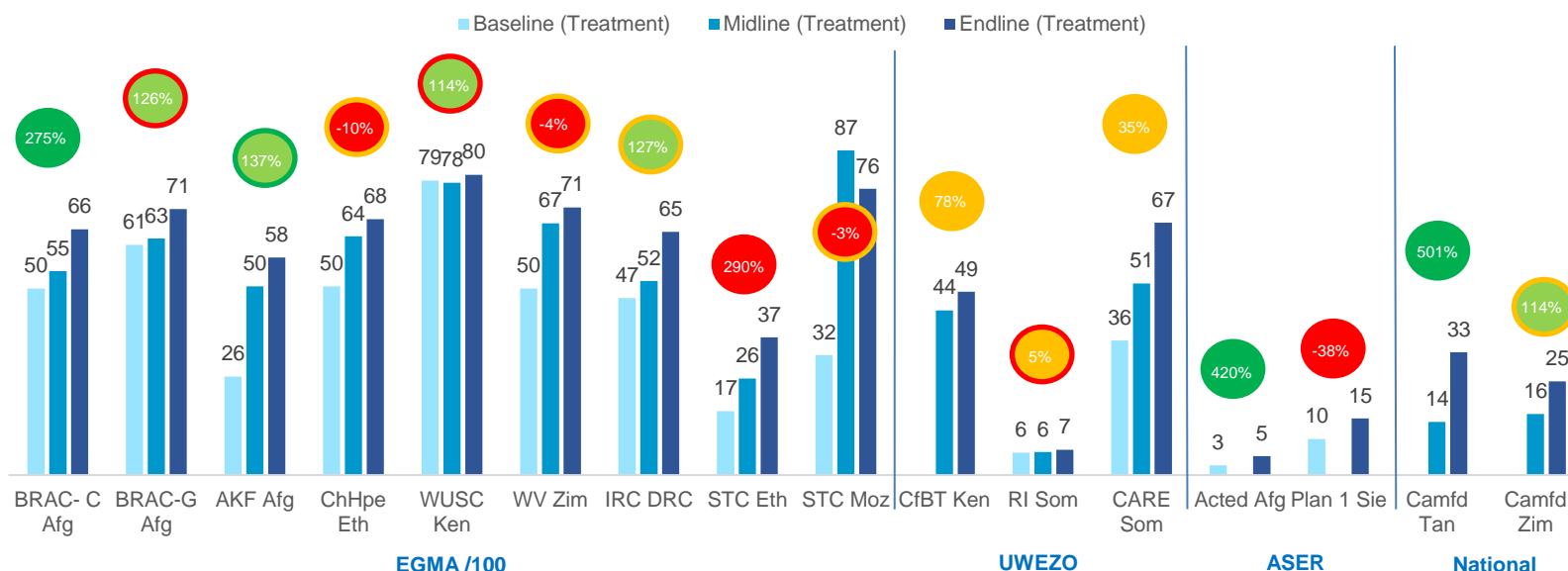
Interpretation

SCW projects' achievements in numeracy were slightly lower than in literacy according to the FM's reanalysis of project data. As for literacy, half of the projects met their endline targets, but only three show very large achievements (compared to five projects' literacy results) and four projects had a negative impact (compared to one project's literacy results). Save the Children (Ethiopia) and Plan (Sierra Leone) have inconclusive data.

WUSC (Kenya) has achieved its endline target despite a very small increase in treatment scores from midline to endline (78% to 80%, from 79% at baseline). This is because of lower benchmark scores at endline than at midline. Indeed, the project had a negative impact at midline: midline treatment scores were lower than benchmarks. Midline treatment scores were then used as endline benchmarks, which explains why endline scores could "easily" exceed them by achieving a marginal increase between midline and endline. A similar conclusion can be drawn for RI (Somalia), which also had a negative midline impact. However, despite smaller benchmarks at endline compared to midline, the project did not manage to achieve its endline target.

Similar to literacy, projects who used EGMA achieved mixed results. However, it is worth noting that the three projects who used UWEZO failed to meet both their midline and endline numeracy targets. Plan (Sierra Leone) and ACTED (Afghanistan) used ASER, a similar test to UWEZO. Only the latter met their target. This suggests that it might be harder to capture relative changes in UWEZO /ASER numeracy scores than in EGRA /EGMA scores over a relatively short timescale (two to three years) because girls may not have time to transition from one test level to the next.

Looking at the overall impact from baseline to endline, only four projects met their target. With the exception of Camfed (Tanzania), three of the projects are located in Afghanistan and do not have control groups. This is much lower than for literacy for which ten projects met or exceeded their targets between baseline and endline.

Conclusion and limitations

> Projects' learning achievements need to be assessed with each project's specific methodology in mind. It is challenging to compare projects who used different evaluation designs (such as those using control groups with those who did not use control groups), as it is to compare two projects who used different tests (although all scores have been standardised). Besides, projects have different ToCs and implemented their activities to different standards, which we assume are similar in this section. With these limitations in mind, our review of SCW projects' achievements suggests that targets have been comparatively easier to reach for some projects than others.

- **In the absence of control groups, the way benchmarks are set has a large influence on projects' learning achievements.**
  In the case of BRAC's (Afghanistan) Community Schools, the benchmark grades for midline and endline achievements are systematically the same as at baseline, based on the assumption that girls' scores would not have progressed at all in the absence of BRAC's intervention. The same can be said of WUSC's (Kenya) benchmarks, of ACTED's (Afghanistan) YDC and Literacy Classes, and of the majority of AKF's (Afghanistan) cohort: for six grade levels out of ten, the midline score is used as the endline benchmark. This might partly explain why projects whose impact was measured against benchmarks show larger achievements than projects whose impact was compared to a control group.

- **When using control groups, the validity of learning achievements depends on the comparability of the treatment and control groups.**
  In the case of Save the Children (Ethiopia), the two groups scored very differently at baseline: the control group's literacy and numeracy scores were both more than twice as large as those of the treatment group. This casts doubt over the validity of the project's impact evaluation design from the start.

- **The size of standard deviations, which form the basis of learning targets, depends on sample sizes and the type of test used.**
  Although standard deviations can be virtually calculated for any type of learning test and score, they are not necessarily adapted to scores expressed in levels, when the number of possible levels is low, such as UWEZO or ASER. Threshold effects, when transitioning from one level to another, can bias standard deviations upwardly or downwardly. Besides, for small samples, standard deviations are a function of sample size. They can therefore be artificially high if calculated from large sample sizes, and low when calculated from small sample sizes. The sample sizes chosen for learning achievements therefore has an important impact on the size of projects' targets.

- **Endline achievements are calculated using midline learning scores as a "starting point".**
  Hence, projects who did not meet their midline targets (whose midline learning scores are relatively low) have a relatively higher chance of achieving their endline target. Conversely, it is more difficult for projects who largely over-achieved their midline target to show a large achievement at endline.

- **Learning achievements do not necessarily reflect the progress of the treatment group's learning scores over time.**
  Although the size of achievement depends on how targets were set, on the type of evaluation design used and on the adjustment calculated by the FM and external evaluators, one would expect a rough comparison between the baseline and endline treatment scores to give a hint about whether the project has achieved their target or not. We do not observe this. For example, World Vision's (Zimbabwe) treatment literacy scores have more than doubled from baseline to endline (33 wpm to 77 wpm) but the project achieved only 18% of their target. CARE's (Somalia) numeracy scores show a similar pattern. Conversely, WUSC (Kenya) met their endline numeracy target, but improved their EGMA scores by only two points. Projects' achievements depend on control and benchmark scores as much as on treatment scores, which underlines the importance of setting benchmarks rigorously and ensuring the validity of quasi-experimental designs.

- **Most of the learning achievements claimed by projects have not been captured by our EM data**. Although the EM sample measure the effects of projects on their wider target communities (and its power is not large enough to capture small changes over the control group), it could have captured some of the largest impacts claimed by projects. This is not the case. The only positive and significant learning impacts observed in the EM data are in AKF's (Afghanistan) target area in numeracy and CfBT's (Kenya) target area in literacy. It is worth noting though, that except for Camfed, all projects who achieved a very large impact (> 150%) in either literacy or in numeracy used benchmarks instead of control groups. It might be that these projects, had they used control groups, may not have achieved as large an impact.

3.3.6 Outcomes by subgroup analysis based on projects' own data

In this sub-section, we explore the trajectories of attendance, literacy and numeracy outcomes on a range of subgroups based on the analysis of SCW projects' own quantitative data. As far as possible, we used baseline, midline and endline datasets and compared the trajectories of treatment and control groups using difference-in-difference estimators. However, this was not always possible because some projects only submitted endline data, or their endline data was not linkable with their baseline and midline data. Besides, six out of fourteen SCW projects do not have a control group, which restricted our analysis to a before-after comparison of treatment areas. Finally, variables' labels were often missing which prevented us to properly identify treatment status, subgroups and test variables. As far as possible, we tried to replicate the analysis from the outcome spreadsheets to ensure our calculations were correct. Further details on the methodology and associated limitations can be found in Annex E. Table 16 shows the data available for each SCW project.

Table 16: Available projects' data for outcome subgroup analysis

Available projects' data	# of proj.	BRAC Afg GOV	BRAC Afg COM	AFK Afg	ACTD Afg	IRC DRC	STC Eth	ChHp Eth	WUSC Ken	CfBT Ken	STC Moz	Plan 1 S-L	RI Som	CARE Som[1]	WV Zim	Camf Tan	Camf Zim
Endline	16	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Midline	12	✓	✓	✓	✓			✓		✓	✓		✓	✓	✓	✓	✓
Baseline	6	✓		✓	✓			✓					✓		✓		
Control group	9					✓	✓	✓		✓	✓	✓			✓	✓	✓
Attendance	5							✓			✓				✓	✓	✓

Note: (1) CARE Somalia data covers primary school girls only.
(2) Two independent sets of data were submitted for BRAC Government Schools and Community Schools (Afghanistan), as well for Camfed in Zimbabwe and Camfed in Tanzania, representing separate cohorts of girls. They will be analysed separately, as different projects. The total number of "projects" considered in this analysis is therefore equal to sixteen.

Subgroups are defined at each wave. For age and grade levels for instance, this means that we compare Lower Primary girls at midline with Lower Primary girls at endline. We only consider subgroups for which data is available from at least three projects. Table 17 shows difference-in-difference indicators, or simple differences, across SCW using the following symbols:

✦	The DiD indicator is positive and statistically significant at the 5% level. Notation: DiD.
✓	The simple difference (before-after) indicator is positive and statistically significant at the 5% level in treatment areas. Notation: SiD.
⊘	Neither the DiD nor the simple difference is positive and statistically significant.
	The cell background is greyed when the DiD / SiD indicators are calculated from midline to endline. It is white when calculated from baseline to endline.

Notes: For projects who have a control group, we first calculate the DiD indicator. If it is statistically significant, we report it. If it is not, we calculate the simple difference and report it if statistically significant. We only show indicators that are based on at least 30 observations. Projects who do not have a control group are filled in light blue in Table 17: BRAC (Afg), AKF (Afg), ACTED (Afg), RI (Som) and CARE (Som). For projects whose baseline data is available, differences are calculated between baseline and endline. Otherwise, differences are calculated between midline and endline. The four projects that only have endline data is available will be described separately.

All girls

We were able to calculate differences in literacy and numeracy scores over time for 12 projects for which we have at least endline and midline data. Out of these twelve projects, six have a control group: ChildHope (Ethiopia), CfBT (Kenya), Save the Children (Mozambique), World Vision (Zimbabwe), Camfed (Zimbabwe) and Camfed (Tanzania). In literacy, only three of these six projects show a positive and statistically significant DiD: the two Camfed projects and CfBT. In numeracy, only Camfed (Tanzania) shows a statistically significant DiD. Save the Children (Mozambique) is the only project showing non-significant difference in literacy and numeracy. However, results for this project are inconclusive because we were not able to replicate the learning analysis from their outcome spreadsheets. More details around replication issues, as well as a table showing results for all subgroups, can be found in Annex I.

School phase

Lower Primary and Upper Primary school girls are the most represented subgroups in our data, which is consistent with projects' targeting. Most projects show at least a significant SiD in learning for these two subgroups, except Save the Children (Mozambique) and RI (Somalia). ChildHope (Ethiopia) shows one for Lower Primary but not for Upper Primary. Conversely, CfBT (Kenya) is the only project showing a significant DiD for both literacy and numeracy for Lower Primary girls. Camfed show a significant DiD between midline and endline for Lower Secondary girls, in both literacy and numeracy in Tanzania, in literacy only in Zimbabwe.

Attendance data has not changed significantly for any projects, except ChildHope (Ethiopia) which shows a significant SiD for Lower Secondary girls.

Age subgroups

SCW projects report learning scores for a range of age categories. Girls from 6 to 17 are covered by almost all of the twelve projects for which we have data. Most of them show a significant SiD in learning over time but very few show significant DiD. None of the age categories show a more widespread improvement than the other. For girls aged 9-11, only ChildHope (Ethiopia) shows significant DiD in literacy and numeracy. CfBT (Kenya) is the only project showing significant DiD for girls aged 12-13, and the only one to show significant DiD for girls aged 14-15 (in literacy only). The two Camfed (Tanzania and Zimbabwe) projects show a significant DiD in literacy for girls aged 16-17, but only Camfed (Tanzania) shows the same for numeracy. Camfed (Tanzania) also exhibits a 5%-DiD in literacy for 18-19-year-old girls, but their effect on the learning of girls older than 19 is not significant.

Schooling subgroups

All projects' samples include in-school girls but only three demonstrate a significant DiD in literacy: Camfed (Tanzania and Zimbabwe) and CfBT (Kenya). Camfed (Tanzania) is the only project showing a significant DiD in numeracy. CfBT (Kenya) also shows a significant DiD in numeracy for girls who never repeated a grade, but fails to show a similar improvement for girls who have repeated a grade. Otherwise, impact on literacy and numeracy is broadly similar. Interestingly, AKF (Afghanistan) shows a positive SiD in numeracy for out-of-school and for girls who dropped out, but fails to exhibit the same increase in literacy.

However, none of the SCW projects are able to demonstrate a positive and significant DiD for out-of-school girls in either literacy or numeracy. This is mostly because the two best performing projects (those showing the most DiD), Camfed (Tanzania and Zimbabwe) and CfBT (Kenya), do not include out-of-school girls in their sample. For other projects, the absence of effects can be due to small samples that do not allow to show statistical significance, even if the size of the difference is relatively large (which points to the inability of projects' evaluation designs and data to demonstrate an effect on out-of-school girls). In RI (Somalia) areas, for instance, out-of-school girls have improved more than in-school girls since baseline in treatment areas, but do not show a significant DiD. Generally however, out-of-school girls' improvements have been comparable to those achieved by in-school girls.

Girls with disability

The evidence is mixed for girls with disabilities. BRAC Community Schools (Afghanistan) demonstrate a significant SiD in literacy and numeracy for girls with no disability, but no significant improvement for girls with a disability. Similarly, Camfed (Zimbabwe) in literacy, as well as CfBT (Kenya) and RI (Somalia) in both literacy and numeracy, demonstrate smaller improvements for girls with disability than for girls with no disability. Conversely, in

Camfed (Tanzania) girls with a disability show a significant DiD in literacy. They even improved more in numeracy in absolute terms (DiD = 12 pp) than girls with no disability (DiD = 6 pp), but their sample sizes were just not large enough to show statistical significance. ChildHope (Ethiopia) also shows encouraging evidence of girls with disability showing larger DiD in literacy, numeracy and attendance, although for the former sample sizes were too small (around 50 girls) to demonstrate statistical significance.

Economic subgroups

Table 17 presents different indicators related to poverty coming from three variables: ability to meet basic needs, hunger or lack of basic resources, and difficulty to afford school. SCW projects do not seem to have had a greater impact on the poorest among their target communities because girls from the marginalised subgroups tend to show the same patterns of change as other girls. It is worth noting that Save the Children (Mozambique) and CfBT (Kenya) show results for the subgroup "Any of the above" but not for "None of the above", because too few respondents chose the latter option. In Save the Children's (Mozambique) target areas, households who said they are unable to meet basic needs show a significant DiD in literacy between midline and endline, and a significant SiD in attendance, as opposed to households who reported being able to meet basic needs. CfBT (Kenya) exhibits an opposite pattern, with girls able to meet basic needs showing significant DiD in learning, and girls for whom schooling is difficult to afford showing a significant DiD in literacy.

Parental characteristics

As with our EM survey, projects' surveys usually do not contain any specific questions on whether the girl is an orphan, but rather ask whether the girl lives with their parents. This can cover a variety of cases such as girls attending boarding school or girls whose parents have migrated to work abroad. It is therefore difficult to conclude anything on the impact of the SCW on such a heterogeneous population. It is worth noting though that ChildHope (Ethiopia) shows a significant DiD for the attendance of girls who live without parents. The project also shows better literacy results for girls whose PCG cannot read or write in the language of instruction (LOI). In a similar way, World Vision (Zimbabwe) shows a significant DiD in the literacy and numeracy of scores of girls whose PCG cannot read or write a letter in LOI, compared to no significant effect on other girls.

Insecurity and journey to school

Evidence overall is inconclusive in terms of insecurity and long distance to school. In ChildHope (Ethiopia), the journey to school takes more than one hour for only 10% of girls. The project did not manage to improve their learning results and even showed negative DiD in both literacy and numeracy. However, the attendance of these girls improved significantly (DiD = 7 pp), which tends to show that the project managed to make them attend school more regularly. However, this effect did not translate into better learning over the course of the programme. World Vision (Zimbabwe), which set up a bicycle scheme, did not manage to significantly improve either the attendance or the learning for girls who have a long journey to school.

Table 17: Subgroup analysis of SCW projects' data

Projects' data	BRAC Afg GOV EGRA wpm (LIT)	BRAC Afg GOV EGMA/100 (NUM)	BRAC Afg COM EGRA wpm (LIT)	BRAC Afg COM EGMA/100 (NUM)	STC Moz EGRA comp. (LIT)	STC Moz EGMA overall (NUM)	STC Moz Att % (ATT)	Camfed Tan Nat. test % (LIT)	Camfed Tan Nat. test % (NUM)	Camfed Tan Att % (ATT)	Camfed Zim Nat. test % (LIT)	Camfed Zim Nat. test % (NUM)	Camfed Zim Att % (ATT)	ChHpe Eth EGRA wpm (LIT)	ChHpe Eth EGMA/100 (NUM)	ChHpe Eth Att % (ATT)	ACTED Afg ASER 4 lev. (LIT)	ACTED Afg ASER 7 lev. (NUM)	WV Zim EGRA wpm (LIT)	WV Zim EGMA/100 (NUM)	WV Zim Att % (ATT)	CfBT Ken UWZ/10 (LIT)	CfBT Ken UWZ % (NUM)	RI Som UWZ/10 (LIT)	RI Som UWZ/10 (NUM)	CARE Som EGRA % (LIT)	CARE Som EGMA/100 (NUM)
All girls	✓	✓	✓	✓	⊘	⊘	⊘	✦	✦	⊘	✦	✓	⊘	✓	✓	⊘	✓	✓	✓	✓	⊘	✦	✓	✓	✓	✓	✓
School phase																											
Lower primary					⊘	⊘	⊘							✓	✦	⊘	✓	✓	✓	✓	⊘	✦	✦	⊘	⊘	✓	✓
Upper primary	✓	✓			⊘	⊘	⊘							⊘	⊘	⊘	✓	✓	✦	✓	⊘	✦	✓	⊘	⊘	✓	✓
Lower secondary								✦	✦	⊘	✦	✓	⊘	⊘	⊘	✓			✓	⊘	⊘						
Upper secondary																											
Age subgroups																											
< 6 years old																											
6 – 8 years old					⊘	⊘	✓							✓	✓		✓	✓	✓	✓	⊘	✓	⊘	⊘	⊘		
9 – 11 years old	✓	✓												✦	✦	⊘	✓	✓	✓	✓	⊘	✓	✓	✓	✓	✓	✓
12 – 13 years old	✓	✓	✓	⊘										⊘	✓	⊘	✓	✓	✦	✦		✓	✓	✓	✓	✓	✓
14 – 15 years old	✓	✓	⊘	✓										⊘	✓	⊘	✓	✓	✦	✓		✓	✓	✓	✓	✓	✓
16 – 17 years old	✓	✓	✓	✓				✦	✦	⊘	✦	✓	⊘	✓	✓	⊘	✓	✓	⊘	⊘	⊘	⊘	⊘	⊘	⊘	⊘	⊘
18 – 19 years old			⊘	⊘				✦	✓	⊘	✓	✓	⊘	⊘	⊘	⊘	✓	✓									
> 19 years old								⊘	⊘	⊘																	
Schooling subgroups																											
In school	✓	✓	✓	✓	⊘	⊘	⊘	✦	✓	⊘	✦	✓	⊘	✓	✓	✓	✓	✓	✓	✓	⊘	✦	✦	✓	✓	✓	✓
Out of school					⊘	⊘								✓	✓		✓	✓	✓	✓	⊘			✓	✓		
Dropped out														⊘	✓												
Never attended																								⊘	⊘		
Has repeated grade					⊘	⊘	⊘							✓	✓	✓						✓	✓	✓	✓		
Has never repeated grade					⊘	⊘	⊘							✓	✓	⊘						✓	✦	✓	✓		
Girls with disability																											
Disabled girl			⊘	⊘				✦	⊘	⊘	✓	✓	⊘	✓	✦	✦			✓	✓	⊘	✓	✓	⊘	⊘		
Not disabled	✓	✓	✓	✓				✦	✦	⊘	✦	✓	⊘	✓	✓	⊘			✓	✓	⊘	✦	✦	✓	✓		
Economic subgroups																											

STEP CHANGE WINDOW ENDLINE EVALUATION REPORT

Projects' data	BRAC Afg GOV EGRA wpm (LIT)	EGMA /100 (NUM)	BRAC Afg COM EGRA wpm (LIT)	EGMA /100 (NUM)	STC Moz EGRA comp. (LIT)	EGMA overall (NUM)	Att % (ATT)	Camfed Tan Nat. test % (LIT)	Nat. test % (NUM)	Att % (ATT)	Camfed Zim Nat. test % (LIT)	Nat. test % (NUM)	Att % (ATT)	ChHpe Eth EGRA wpm (LIT)	EGMA /100 (NUM)	Att % (ATT)	ACTED Afg ASER 4 lev. (LIT)	ASER 7 lev. (NUM)	WV Zim EGRA wpm (LIT)	EGMA /100 (NUM)	Att % (ATT)	CfBT Ken UWZ /10 (LIT)	UWZ % (NUM)	RI Som UWZ /10 (LIT)	UWZ /10 (NUM)	CARE Som EGRA % (LIT)	EGMA /100 (NUM)
Unable to meet basic needs					✦	⊘	✔							✔	✔	⊘						✔	✔				
Able to meet basic needs					⊘	⊘	⊘							⊘	✔	⊘						✦	✦				
Hunger			⊘	✔	⊘	⊘	⊘	⊘	✔	⊘	✔	✔	⊘	✔	✔	⊘			✔	✔	⊘	✔	✦				
Not enough clean water			⊘	✔	⊘	⊘								✔	✔	⊘						✦	✦				
Not enough medicines			⊘	✔	⊘	⊘								✔	✔	⊘						✔	✔				
Not enough cash income			⊘	✔	⊘	⊘								✔	✔	⊘						✔	✔				
Any of the three above			✔	✔	⊘	⊘	⊘							✔	✔	⊘						✦	✦				
None of the three above			✔	✔										✔	✔	✔											
Difficult to afford school					⊘	⊘	⊘							✔	✔	⊘						✔	✔	✔	✔		
Not difficult to afford school					⊘	⊘	⊘							✔	✔	✔						✦	✔	✔	✔		
Parental characteristics																											
Lives without parents					⊘	⊘	⊘	✦	✦	⊘	✔	✔	⊘	⊘	✔	✦			✔	✔	⊘					✔	✔
Lives without mother					⊘	⊘	⊘	✦	✦	✔	✔	✔	⊘	⊘	✔	⊘			✔	✔	⊘						
Lives without father					⊘	⊘	⊘	✦	✦	⊘	✦	✔	⊘	✔	✔	⊘			✔	✔	⊘					✔	✔
Lives with both parents			✔	✔	⊘	⊘	⊘	✦	✦	⊘	✦	✔	⊘	✔	✦	⊘			✔	✔	⊘			✔	✔	✔	✔
PCG can read/ write LOI					⊘	⊘	⊘	✦	✦	⊘	✔	✔	⊘	⊘	✔	✔			✔	✔	⊘			✔	✔	✔	✔
PCG cannot read/ write LOI			✔	✔	⊘	⊘	⊘	✦	✦	⊘	✦	✔	⊘	✔	✔	⊘			✦	✦	⊘			✔	⊘	✔	✔
Insecurity / Journey to school																											
Unsafe area			⊘	✔	⊘	⊘	⊘							⊘	⊘	⊘						✔	⊘	✔	✔		
Safe area			✔	✔	⊘	⊘	✔							✔	✔	⊘						✦	✦	✔	✔		
Journey to school > 1h								✦	⊘	⊘	✦	✔	⊘	⊘	✔	✦			✔	✔	⊘						
Journey to school < 1h								✦	✦	✦	✔	✔	⊘	✔	✔	✔			✔	✔	⊘						

Key findings – What impact has the GEC had on enabling marginalised girls to be in school and learn?

Being-in-school outcomes

- The EM data showed that in-school enrolment increased slightly among SCW target communities, from 76% at baseline to 82% at endline. However, the increase has been similar in control areas.

- The proportion of out-of-school girls among target communities is much higher in some contexts (Somalia, Ethiopian Afar region, Kenya refugee camps). Further analysis shows that in these contexts and in the Afghanistan projects' areas, most out-of-school girls have never attended schools, while in other contexts, out-of-school girls are older girls who dropped out after completing primary school.

- The EM evidence suggests that SCW projects have managed to slightly increase the retention rate of older girls (aged 14-15), which is an encouraging finding. However, most girls are still lagging behind in terms of the school phase they are in: at endline, more than half of the girls aged 14-15 are still enrolled in primary school.

- Gender comparisons show that among SCW communities, girls now have higher enrolment rates than boys across all age categories. But trends are similar across treatment and control areas, which implies that they may not be an effect of GEC intervention.

- Attendance varies very little across time and across project areas. This is mostly due to the inability of our household-based measures of attendance to pick up small variations in attendance rates. However, the school-based measure introduced at midline shows similar patterns, which suggests that attendance has remained constant overall among SCW communities.

Learning outcomes

- Literacy and numeracy levels have steadily improved among girls in SCW communities. On average, girls improved by 16 words-per-minute and by 25 percentage points (EGMA/100). However, we observe no significant differences between treatment and control communities in the EM data.

- Window-level findings hide disparities across project areas and subgroups. Our EM findings show that older girls have realised the largest learning gains in absolute terms. However, further analysis suggests that lower-performing girls have improved more than other girls of the same age. Out-of-school girls, for instance, have improved by 20 wpm in literacy and by 34 pp in numeracy.

- SCW projects do not seem to have had a specific effect on subgroups, according to projects' data analysis. Subgroups usually show similar patterns of changes as the rest of the population. Our analysis has however been limited by the small, often very small, sample sizes associated to some subgroups, and to the varying degrees of available and exploitable data across projects.

- Girls are still lagging largely behind international benchmarks of oral reading: the basic comprehension level of 45 wpm, which should be attained during the second year of primary, is only attained by girls over the course of their sixth year of school.

- Our analysis of EGRA subtasks shows that girls' scores have improved the most in the easiest literacy subtasks compared to control areas. The first EGRA subtask, invented word, is the only subtask showing a significant positive difference-in-difference. The two most difficult subtasks even show a negative difference-in-difference compared to control areas. We observe a different pattern for EGMA: girls have improved the most in intermediary subtasks, addition and subtraction.

- The progress in EGRA wpm score seems to be correlated more to the time spent in school than the EGMA total score: there are large differences between the literacy scores of out-of-school girls, enrolled girls with low attendance, and enrolled girls with high attendance, but relatively small differences between their numeracy scores.

- Our analysis of EM data suggests that some activities have had a significant positive impact on girls' learning, such as the provision of scholarships or bursaries, study groups, special tutoring or help with school work, as well as teachers' training in general teaching methods and mentoring. These activities

have an effect on both literacy and numeracy scores, and a similar one across treatment and control areas. Conversely, talks about enrolling or staying in school, provision of additional funds, school book provision as well as teacher training in gender sensitive pedagogy and special needs education, do not seem to have directly led to significant learning gains over the course of the programme.

Key lessons learned

- Findings from the EM data suggest that a substantial proportion of girls among target communities are still out-of-school. In some contexts, these girls have never been enrolled or have attended just a couple of years of school. In other contexts, girls are dropping out as they grow older and reach the end of primary. This suggests that the **population of out-of-school girls is not homogeneous across the SCW**. While in some contexts, encouraging retention and helping transition to secondary school may be sufficient to ensure that girls keep learning at the right pace, in other contexts where out-of-school girls are excluded from the school system, other types of intervention are needed that directly aim at improving learning, for instance by creating learning environments which can act as substitutes for gaps in their schooling.

- The **measurement of in-school attendance over the course of the GEC has proved challenging**. Similar constraints were experienced in the EM and projects' research. Household-based measures, usually asking girls' primary caregiver, are not precise enough to capture changes over time. School-based measures are highly dependent on the reliability of school records and registers, which have often been found to be of very poor quality. Some projects have implemented innovative mechanisms and protocols that allow more rigorous measures of attendance in the schools they are working in. But these require frequent measurement and regular follow-ups in the field. Attendance can more reliably and accurately be captured as part of projects' monitoring data, but does not really lend itself to the one-off and distant measurements of an impact evaluation.

- At endline, **girls are still lagging far behind literacy benchmarks**. The number of wpm girls can read towards the end of primary school (during their sixth year of school) should have been reached throughout the course of the second year of primary, according to international benchmarks. This means that **girls spend most of their primary school years learning to read rather than reading to learn**. The ability to read opens the door to other subjects so this lag may not only impact girls' literacy but their learning progress overall.

- Findings from our EM data suggest that **teachers' training in general teaching pedagogy, as well as** extra-curricular activities aimed at directly improving learning such as **special tutoring or help with schoolwork, are among the most effective interventions to improve girls' literacy and numeracy** in both treatment and control areas. In some contexts, where current learning environments in school are not sufficient to make girls improve at a steady pace, long-term investments in these types of intervention may be required. The effectiveness of the different SCW activities are explored further in Section 3.4.

- Our analysis of the EM data suggests that SCW projects did not improve literacy and numeracy scores more than in control areas, except for the subtasks of EGRA invented words and EGMA level 1 addition. To ensure that girls fully learn, projects will need to ensure that **all dimensions of literacy and numeracy are adequately addressed, and that girls from all parts of the learning spectrum increase their learning**. This may require tracking the outcomes of girls from the lower end of the learning distributions, as well as of girls belonging to marginalised subgroups of the population, and ensure they are allowed sufficient sizes in projects' samples.

- The **absence of an overall effect of SCW projects on literacy and numeracy is consistent with our exposure findings** from Section 3.2, which show that target communities have been more exposed to education-related activities at endline than at baseline, but to a similar extent in treatment and control areas. A similar conclusion can be drawn from projects' achievements (based on their own data): in literacy, only three projects whose evaluation design includes a control group reached their baseline to endline target in literacy (IRC in DRC, CfBT in Kenya and Camfed in Zimbabwe and Tanzania), while

only Camfed reached their target in numeracy. This is likely due to a mix of factors including: the insufficient size and /or magnitude of impact achieved by SCW projects on their target communities; overcrowding of project areas with GEC and non-GEC actors; possibility of positive spillover effects; and poor sampling from the onset – the lists of sampling points provided by projects before baseline may not accurately reflect their areas of intervention any more, and some sampling points may have been too close together to prevent large spill over effects from treatment to control areas (see Annex F). These factors show the **difficulty of implementing large and rigorous quasi-experimental designs in such contexts and the need to improve projects' and external evaluators' M&E capacity** overall.

3.4 What has worked, why and with what effects?

SCW projects aim to address a range of barriers to girls being in school and learning effectively. We undertake intervention mapping (Annex B) to identify the reported effectiveness of project interventions on the relevant barriers identified in their areas. This is triangulated with primary data gathered by projects themselves and by the EM. Unless otherwise stated, all data presented in the main body of the text is harvested from projects' endline evaluation reports. A detailed presentation of the methodology is presented below. For summary tables for each intervention and each barrier, refer to Annex J.

Measuring changes in barriers to girls' education and assessing the effectiveness of interventions

To further develop their interventions, projects identified specific barriers at baseline that they assume drive educational marginalisation in their target areas. While some of these barriers are structural or otherwise beyond projects' direct control (such as the occurrence of droughts), others may be tackled through targeted interventions and support (such as a lack of adequate sanitation facilities in schools).

This report presents evidence of the **most reported barriers** *perceived* by projects to be preventing girls from attending school and learning. As such, barriers may not be *actual* barriers but the influence of these **barriers, either actual or perceived**, which is assumed here to prevent girls from attending school and learning.

Following a similar approach to that used for Innovation Window projects, barriers were categorised across the key thematic areas (refer to Table 23 in Annex J) that emerged from projects' endline reports. We follow a four-staged approach to assessing the most and least prevalent barriers and the effectiveness of interventions.

1. **Identify barriers**. The metrics used to assess the prevalence of barriers are derived from whether the reported barriers are deemed as prevalent or not prevalent by the projects in their reports.

   Across the SCW and for each of the identified barriers, we discuss the **number of projects which have reported the existence of the specified barrier** in their target areas. The ranking of reported barriers (from most to least reported) gives the relative prevalence of barriers compared to others across SCW projects[64].

2. **Assess changes in barriers**. The second stage involves a **project-by-project assessment of findings** in order to assess whether the evidence was present, absent, inconclusive or not reported by projects for the specified barriers that projects' activities aimed to tackle.

   Barriers have either lessened, not changed or worsened since baseline, or evidence is inconclusive. The *extent* to which barriers have changed could not always be assessed. Therefore, changes in barriers should be interpreted more as a '**direction of change**', either positive, neutral or negative.

3. **Link changes in barriers to interventions**. Finally, the third stage relates to assessing the **effectiveness of interventions** from projects' findings, e.g. whether the interventions are deemed effective or not by the projects, and whether sufficient and conclusive evidence has been presented to support this.

---

[64] It is important to note that the data collected by projects is focused on their **target groups** rather than the general population or communities in which their target groups live. This means that unless projects have undertaken a population study as part of their baseline, midline and endline research, those barriers that are most reported may not necessarily be the most prevalent in the communities in which they are working.

We indicate the origin of the findings by referring to individual Project Endline Evaluation Reports and we indicate our reservations wherever projects themselves have expressed reservations, or where these findings are not consistent with those reported by the EM or by projects in their primary data (see below).

Findings are then triangulated with primary data gathered by: (i) the EM; and (ii) projects themselves. For a further discussion of the methodology and limitations involved in data collection, please refer to Section 2).

It is important to note that for Plan (Sierra Leone), the outbreak of Ebola significantly impacted project delivery between midline and endline. In addition, activities in ACTED (Afghanistan) areas stopped at midline following a decision by the FM based on the performance of the project. As such, findings presented date back to their midline report submitted in 2016[65].

What effects has the SCW had on barriers to girls' education and with what impact on learning and being in school?

In this section, we report on the following barriers with regards to their impact on girls' education across the SCW:

- Poverty and livelihoods;
- School-related factors;
- Attitudes towards girls' education;
- Girls' aspirations; and
- Violence and safety.

Poverty and livelihoods

All SCW projects address poverty and livelihood issues as key barriers to girls' education at endline, such as the affordability of school expenses, household chores, and the level of household income, all of which can make it difficult for girls to attend school regularly and find time to study.

At endline, the most common reported improvements were in the cost of schooling (five projects) and the availability of educational resources at home (three projects), although certain barriers prevail at endline. One project reported that the supply of educational resources at home had got worse, with four projects reporting a worsening in the amount of housework commitments of girls, reflecting additional responsibilities for girls as they age.

Key findings

Scholarships and bursaries

- Scholarships, bursaries and stipends are directly helping keep girls in school by ensuring timely payment of school fees and increasing girls' motivation.
- There is some encouraging evidence that bursaries and scholarships might be indirectly linked to improved learning outcomes by improving attendance.
- There is concern about increasing schooling costs and the impact on households' ability to afford educational costs in the future.

Housework commitments

---

[65] Girls' Education Challenge (2017). Project Completion Review – ACTED (Afghanistan), p.7.

- Projects have not been effective in reducing the amount of time spent on housework commitments (for example through community sensitisation activities), with girls still having to juggle their school and housework commitments.

- Household chores continue to be borne by girls, with little evidence of attitudinal change towards girls' household responsibilities in terms of the equitable distribution of tasks or time available for study.

In-kind support

- Providing sanitary wear and menstrual supplies is helping to encourage more regular attendance at school.

Loans, savings and income-generating activities

- Interventions facilitating access to loans and savings and supporting income-generation are encouraging caregivers to spend more on girls' education

Cost of schooling

There is some evidence that providing scholarships helps keep girls in school through timely payment of school fees. According to IRC's (DRC) endline report, attendance rates increased overall from baseline to endline, and to a greater extent in intervention schools. Qualitative evidence suggests girls were no longer chased away from school due to the non-payment of school fees, meaning that they were no longer dipping in and out of class pending payment of their fees. These 'chasing away' and 'revolving door' effects were observed in the Girls' Education Challenge Thematic Research Report (2016: 18)[66], which identified that girls' education was repeatedly interrupted as a result of parents' failure to pay direct costs, although scholarships are in part helping to mitigate this through regular payment of school fees.

Similarly, evidence from WUSC (Kenya) suggests that scholarships can act as a powerful incentive to keep girls in school. In WUSC's target areas, a majority of girls receiving scholarships had successfully transitioned to the next year and were continuing their studies, with the retention rate up from 86 per cent at midline to 90 per cent at endline. Qualitative evidence from Head Teachers suggests that this is because scholarships can act as a pull factor to 'lure' parents when deciding whether to send their daughters to schools, as well as a motivating factor for girls themselves, helping to encourage their regular school attendance.

There is some encouraging evidence that bursaries and stipends are helping to offset the costs of education, with positive effects on enrolment. BRAC's (Afghanistan) project provided stipends to government school girls to encourage re-enrolment and retention of girls who dropped out in the past, with a majority of girls suggesting that their families could otherwise not afford the school costs associated with continuing their education. In STC's (Mozambique) project, enrolment rates for girls who received a bursary from the project is six per cent higher at endline compared to girls who did not receive a bursary.

Providing bursaries, stipends and scholarships may be linked to positive learning outcomes in certain contexts, although evidence is limited on why this may be the case. In Camfed's (Zimbabwe and Tanzania) project, financial support (whether through bursaries or the *Step Up Fund[67]*) was linked to increased attendance, while a statistical association was found between those in receipt of bursaries and higher attainment in English over time. The endline report suggests that neither of these interventions were designed to improve learning outcomes, however, other than by enabling girls to remain school and therefore performing better than peers who had dropped out. Qualitative findings from their endline report echoes this, and suggests that the provision of the financial support helped facilitate girls' retention.

Similarly, qualitative evidence from BRAC's (Afghanistan) evaluation report found that girls were using their stipends to cover school-related costs, which could explain larger improvements reported in stipend girls' numeracy and literacy scores from midline to endline compared to non-stipend holders. In addition, at midline, EM qualitative

---

[66] Evaluation Manager Girls' Education Challenge Fund (2016). Narrow windows and Revolving Doors. GEC Thematic Research Report, March.
[67] Adapted from Camfed's Endline Evaluation Report (p122): this is a low cost needs-based financing mechanism offered to girls identified as marginalised within their school or community, according to the definition of marginalisation included in their report.

data found that scholarships were the most commonly mentioned programme interventions by respondents in IRC (DRC), with evidence that they were encouraging girls to be more committed to their studies.

There is concern for the sustainability of scholarships in light of rising tuition fees relative to average household incomes. Both WUSC (Kenya) and CARE (Somalia) raised concerns over the sustainability of these interventions, with WUSC acknowledging that demand for scholarships far exceeded supply. In CARE's project areas, it is estimated that a third of the average per capita income is spent on the costs associated with attending school. Although the proportion of girls receiving scholarships or having their fees waived significantly increased from midline to endline (from 16 per cent to 21 per cent), tuition expenses more than doubled during this period for households paying tuition.

Box 6: Spending on girls' education – what has changed overall?

Findings presented below are based on primary data gathered by projects at endline. Projects presented below are those that at a minimum collected data for treatment areas from midline to endline.
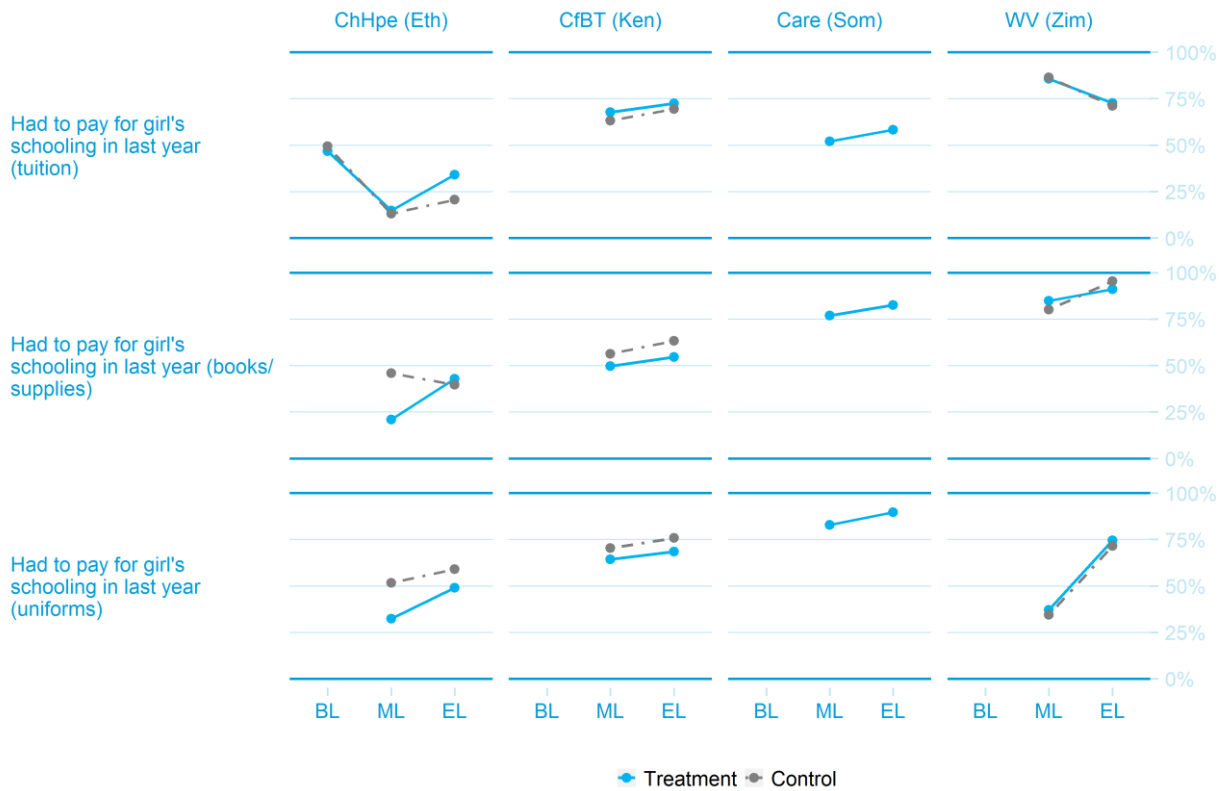
Project data from CfBT (Kenya), CARE (Somalia), Child Hope (Ethiopia) and WV (Zimbabwe), all found an increase in the number of households having to pay for aspects of girls' schooling, including uniforms and school supplies.

However, while an increasing number of households reported having to pay for girls' tuition in Child Hope's, CfBT's and CARE's projects since midline, this trend was in decline for WV. This could echo findings presented in their endline report, which suggested that village savings and loan associations were helping to increase the financial capacity of caregivers and their ability to pay school fees.

Although CfBT provided uniforms to girls as part of back-to-school kits (see further *In-kind support* below)[68], the fact that more households reported having to spend money on uniforms since midline could be testament to their reported popularity and insufficient supply to meet demand. In their endline report, CfBT also cautioned against the lack of sustainability of these interventions, because no uptake mechanism beyond the project's support has been identified at project end.

Housework commitments

Over half of all projects sought to address the time girls spend doing household chores, including cooking and taking care of younger siblings and sick family members. These were predominantly addressed in two ways, including by:

- **raising awareness of housework commitments and its impact on girls,** such as through radio programmes addressing issues related to heavy domestic workloads on girls' education (WUSC (Kenya)), and Mothers' forums (BRAC (Afghanistan)); and

- **intervening directly with parents** to ensure girls attend school (Child Hope (Ethiopia)) and *shuras*[69] (AKF (Afghanistan)).

At endline we see that **projects have generally not been effective in reducing the amount of time spent on housework,** with four projects reporting that girls' housework commitments have worsened since midline, and three projects finding that this barrier to attendance still largely prevails. This echoes our earlier observations on attendance based on EM data, which finds that attendance has more or less remained constant since baseline, with girls on average missing around one day of school a week (see sub-section 3.2.2). This is not surprising, given that there is considerable evidence that girls' domestic responsibilities increase as they age (see, for example, Crivello (2017))[70].

Box 7: Poverty barriers to girls' education – what has changed overall?

Findings presented below are based on primary data gathered by the EM at endline. Treatment averages are shown on top and control averages are shown underneath. Statistically significant changes are indicated in the text below.

Difficult to afford for girl to go to school

Time doing non-school work

---
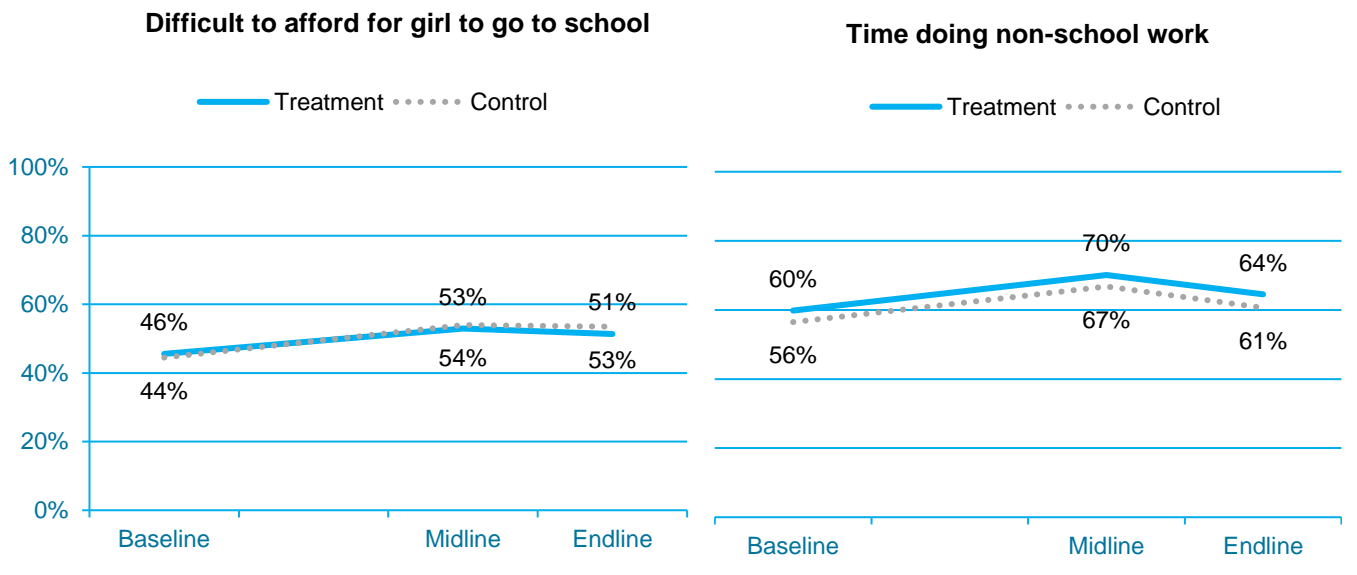
[68] According to their endline evaluation report, CfBT (Kenya) distributed 15,770 back-to-school kits (containing uniforms) against an endline target of 14,000 kits. Of these, 13,778 were distributed to girls. The total number of direct learning beneficiaries (girls) was 88,921.
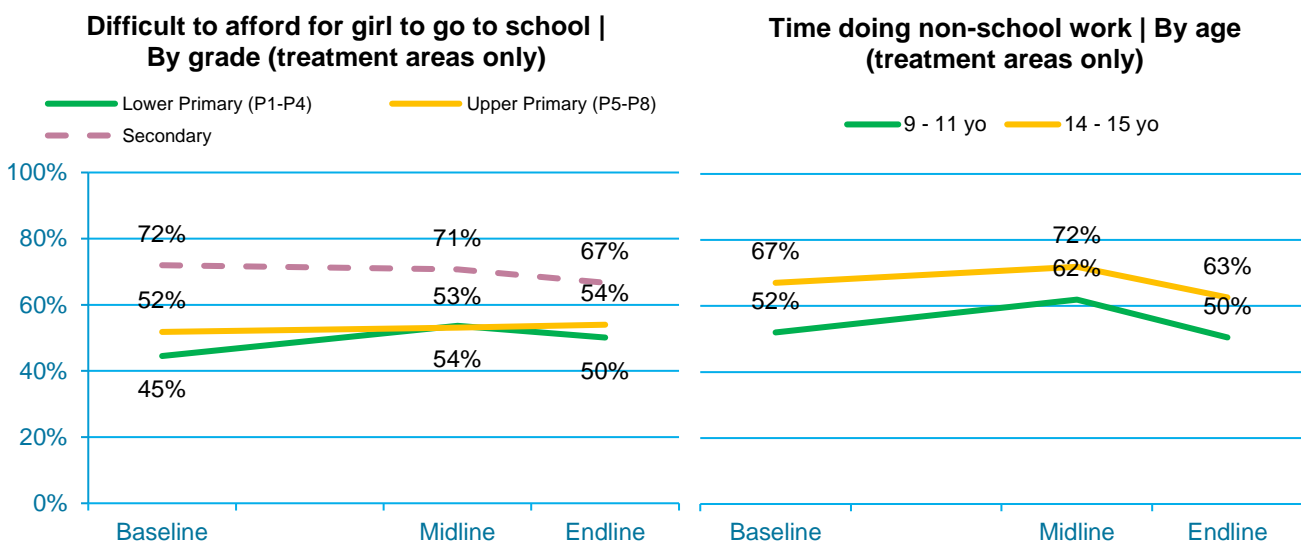[69] These are school management councils (Endline Evaluation Report, AKF: 5).
[70] Crivello, G. (2017). Tracing the links between girls' unpaid care work and women's economic empowerment. Available at: https://www.younglives.org.uk/content/tracing-links-between-girls%E2%80%99-unpaid-care-work-and-women%E2%80%99s-economic-empowerment

Time spent doing non-school work

Data gathered by the EM reveals an increase in the proportion of girls spending time doing non-school work since baseline at the window level. Analysing the data by age (see further below), this confirms the fact that **as girls get older they are required to take on additional chores and housework,** with a higher proportion of adolescent girls reportedly undertaking work outside of school compared to younger girls. This suggests that the increase in the proportion of girls spending time doing non-school work could be attributed to a **cohort effect**, with the same girls of caregivers interviewed at baseline now being required to conduct more chores as a result of reaching adolescence.

Difficult to afford for girl to go to school | By grade (treatment areas only)

Legend: Lower Primary (P1-P4) — Upper Primary (P5-P8) — Secondary

	Baseline	Midline	Endline
Secondary	72%	71%	67%
Upper Primary	52%	53%	54%
Lower Primary	45%	54%	50%

Time doing non-school work | By age (treatment areas only)

Legend: 9 - 11 yo — 14 - 15 yo

	Baseline	Midline	Endline
14 - 15 yo	67%	72%	63%
9 - 11 yo	52%	62%	50%

Difficulty affording sending girls to school

Households are finding it increasingly difficult to afford sending their girls to school, with a higher proportion of caregivers indicating that it is more difficult to afford sending the girl to school at endline than at baseline (51% / 46%, 53% / 44%, -4 pp).

Endline evidence from project evaluation reports suggests that this trend could be **due to the increased costs of schooling relative to households' income, rather than the costs per se.** Plan (Sierra Leone), for example, suggests that the poorest families are increasingly struggling to make ends meet, while WUSC (Kenya) reports a general downward trend in households being able to meet more than their basic needs, which it attributes to a worsening situation among refugee households.

Looking into analysis by sub-group, there are some interesting **variations by grade.** While the affordability of school has remained fairly constant for caregivers of Upper Primary School girls, there are variations among caregivers of Lower Primary and Secondary School girls. At the Lower Primary level, a higher proportion of caregivers in treatment areas at endline report difficulty in being able to afford sending their girls to school (50 per cent) as compared with baseline (45 per cent), whereas an opposite trend is observed at the Secondary level, with fewer caregivers finding it difficult at endline (67 per cent) as compared with baseline (72 per cent).

In SCW countries, girls attending secondary schools are required to pay school fees. So, this finding most likely reflects the fact that economic interventions (such as scholarships) have predominantly targeted secondary school girls and are helping to offset the costs of education for those girls in particular. Although primary schools are officially free across all SCW countries[71], the increase in caregivers finding it difficult to send their girls to

---

[71] According to official government policies, primary school is free in all GEC SCW countries. However, in practice, primary schools in some countries continue to ask students to pay fees (GEC SCW Midline Evaluation Report, 2017: 24).

school could reflect caregivers struggling to meet the costs associated with other 'informal' or indirect schooling expenses (such as books, lunch, uniforms), particularly where they have not been targeted by projects' economic interventions.

Poverty barriers to girls' education – evidence from projects' data

Findings presented below are based on primary data gathered by projects at endline. Projects presented below are those that at a minimum collected data for treatment areas from midline to endline.
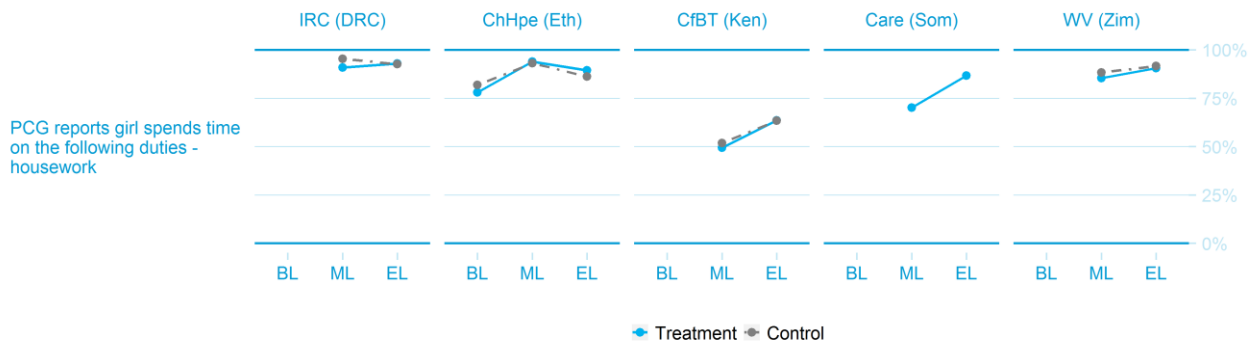
Project data for IRC (DRC), Child Hope (Ethiopia), CfBT (Kenya), CARE (Somalia) and WV (Zimbabwe) all found an increase in the reported number of girls in treatment areas spending time doing housework since midline.
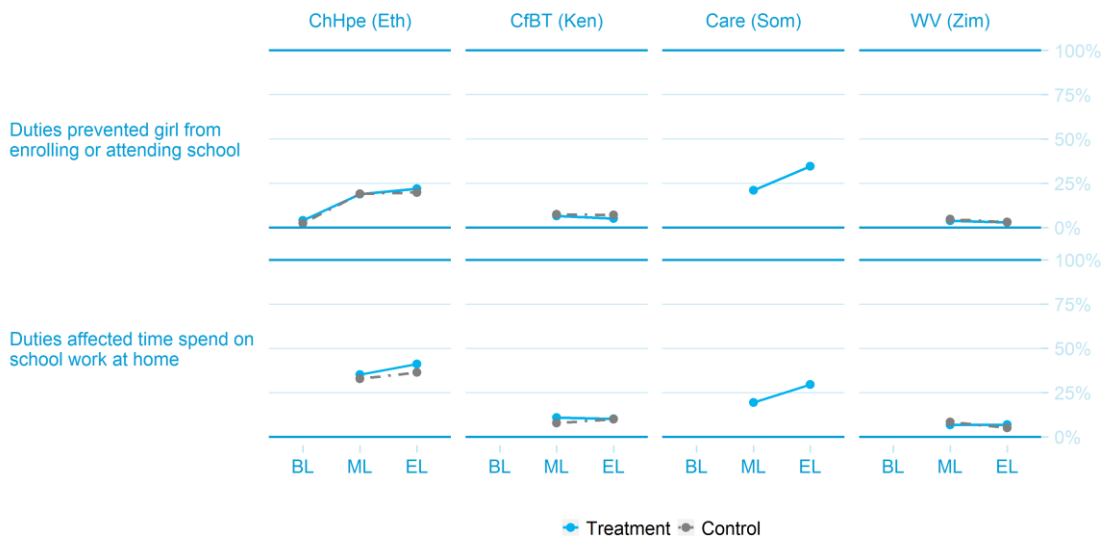
Nevertheless, despite increasing numbers of girls reported to be spending time doing housework at endline, there have not been significant differences in the effect of these duties on girls' reported ability to attend school or complete their homework. This suggests that increasingly girls are being required to juggle their time in order to complete their duties and schoolwork effectively (see further below).

Evidence suggests that girls are having to stretch their time in order to juggle housework and school commitments. Qualitative evidence at endline presented by Child Hope (Ethiopia) showed that girls were able to combine their domestic tasks with schooling, even though the average hours the girls spent at work (domestic and other) significantly increased between midline and endline for both treatment and control groups. This was also the case in BRAC's (Afghanistan) project, with girls suggesting that they were able to organise their time effectively to conduct household labour, while also being able to attend school and do their homework. This is further evidenced by the fact that similar proportions of girls claim to never miss school to help with the management of the house between midline (76 per cent) and endline (78 per cent).

In AKF's (Afghanistan) project areas, there was evidence of school *shuras* playing a key role in supporting girls' attendance by encouraging household members to reduce the pressure on child labour:

> *I try to be present at school and study my lesson. This is clear that it affects our studies and marks. And the chores I do at home is not lessened, but I do them all when I am not at school. I am less absent because of household chores. Because if I am absent then the teacher will tell that to shura members and they tell my parents to send your daughter to school.*

Girl Student, Adraskan District, Community-Based Education class
Endline Evaluation Report, AKF (Afghanistan)

This suggests that *shuras* are effective in encouraging attendance, 'deterring' the girls from skipping school to avoid having to explain their absence to their parents. Similarly, evidence from CfBT (Kenya) indicates that activities such as community conversations and Community Healthcare Worker visits have not been especially effective, with the percentage of households allocating girls with chores increasing steadily from baseline to endline.

Household chores continue to be borne by girls, with little evidence of attitudinal change towards girls' household responsibilities. According to their endline evaluation report, girls in CARE's (Somalia) project areas are still expected to carry out a disproportionate amount of household duties relative to their brothers, meaning they are often late to school. Similarly, Child Hope (Ethiopia) found that there was little change in community attitudes towards household responsibilities, in that they are predominantly carried out by girls, irrespective of whether or not they are in school and regardless of the availability of boys in the household.

There is some evidence that increasing housework commitments may not be as detrimental to learning outcomes as anticipated. Both AKF (Afghanistan) and BRAC (Afghanistan) presented evidence showing significant reading fluency improvements between midline and endline, irrespective of whether or not the girls carry out non-school work. However, this only appears to be true for girls reported to be spending less than a day doing such activities, which suggests that there is a 'critical point' beyond which girls' abilities to juggle their time is no longer sufficient to compensate in learning (see Orkin, 2012[72]). A similar pattern is observed for numeracy outcomes, with girls in both projects showing improvements from midline to endline regardless of how much time they spend working. Again, these improvements cease to be significant for girls spending more than half a day working.

In-kind support

A number of projects identified the lack of educational resources at home as a barrier to girls' attendance and learning outcomes, which projects aimed to address through the provision of in-kind support, such as solar lamps (e.g. WUSC (Kenya), RI (Somalia)), back-to-school kits (Plan (Sierra Leone), CfBT (Kenya)) and menstrual supplies (e.g. CARE (Somalia)), WUSC (Kenya), WV (Zimbabwe), RI (Somalia)).

Providing sanitary wear and menstrual supplies is helping to encourage more regular attendance at school. At endline, RI (Somalia) presented statistical evidence attributing 12 per cent of the variation in girls' attendance to the provision of sanitary kits. WUSC (Kenya), CfBT (Kenya) — and to a lesser extent WV (Zimbabwe) — all presented evidence linking the provision of sanitary wear to higher attendance. Qualitative evidence from across these projects suggests that this is because sanitary wear enables girls to manage their periods more 'discreetly', without needing to skip school for fear of being teased. In this respect, the sanitary wear improved girls' confidence and enabled them to attend school more regularly (see for example Coast et al, 2017[73]).

Three projects directly provided uniforms to the girls: WUSC (Kenya), CfBT (Kenya) and RI (Somalia). At endline, CfBT reported that the attendance of girls who received back-to-school kits (which included uniforms) had increased from midline, with attendance consistently higher than the attendance of girls in intervention schools who did not receive them. Qualitative evidence also shows that the school uniform had made attendance more regular

---

[72] Orkin, K. 2012. Are Work and Schooling Complementary or Competitive for Children in Rural Ethiopia? In Bourdillon, M. and Boyden, J. (eds.), Children Growing Up in Poverty: Multidisciplinary Approaches (Basingstoke: Palgrave): 298-315.
[73] Ernestina Coast, Elizabeth Presler-Marshall and Samantha R. Lattof (2017). An agenda for policy and action to support girls through puberty and menarche. GAGE Digest, ODI, UK. Available at: http://www.gage.odi.org/sites/default/files/2017-07/P%20and%20M%20Digest%20-%20FINAL.pdf

because girls felt more comfortable and were no longer chased away for not being in possession of a school uniform. This chimes with qualitative evidence from RI (Somalia), which suggests that the provision of uniforms helped increase enrolment by lowering the costs of education, thereby enabling parents to afford to send their girls to school.

> *I did not have a school uniform before BTSK. I used to be sent home for uniform. Since I got a uniform I am comfortable and no longer sent home for the same. Uniform makes us equal and are able to come to school regularly and to participate in class activities. School shoes help us because toilets are dirty. Now I come to school every day, and I am grateful.*

> Girl student, County 61
> Endline Evaluation Report, CfBT (Kenya)

Nevertheless, in their endline evaluation report, WUSC (Kenya) reports that there were some concerns of the uniforms as 'ill-fitting, poorly crafted [and] religiously inappropriate' (Endline Evaluation Report, WUSC (Kenya)). Girls believed the uniforms did not foster self-confidence, which could deter girls from attending school regularly, although this cannot be confirmed by scrutiny of attendance data since this was not collected regularly.

In addition to uniforms, these projects provided solar lamps to households to increase the available study time for girls in the evening. **At endline, there is evidence that the lamps are not always being used as intended**, with reports from WUSC (Kenya), RI (Somalia) and CfBT (Kenya) of the lamps being used by beneficiaries other than the targeted girls or in some cases not used by the girls at all (WUSC (Kenya)).

By and large very few projects presented conclusive evidence about the impact of solar lamps on learning outcomes, although RI (Somalia) found a significant association between receiving a solar lamp and higher literacy scores for girls in rural areas affected by drought and conflict. Girls who received solar lamps did slightly better than peers who did not receive lamps, with qualitative evidence attributing this to the fact that girls could now complete their homework in the evening. There is some experimental evidence supporting the idea that solar lamps have limited impact on learning outcomes, in part due to the poor quality of the products (Furukawa, 2014; Shonchoy et al, 2017)[74].

Loans and savings

Interventions such as loans and savings are designed to increase the overall household income to encourage caregivers to invest in girls' education (by paying school fees and associated costs). Examples of interventions used by projects are included below:

- **Village savings and loan associations (VSLAs)** – set up by IRC (DRC), STC (Ethiopia) and WV (Zimbabwe) in the community to provide savings and credit for caregivers. These were accompanied by entrepreneurship and finance training in some cases (IRC and WV) as well as awareness-raising activities on the importance of prioritising household expenditure on girls' education (IRC and STC (Ethiopia)).

- **Livelihood support programmes** – such as the Community Livelihood Fund put in place by STC (Ethiopia) and the Livelihood support programme implemented by Child Hope (Ethiopia). This was intended to support families with insecure livelihoods, and worked by providing assets to caregivers to encourage them to set up their own enterprises and businesses.

- **Cash transfers** – which was implemented by Child Hope (Ethiopia). This involved providing cash transfers to the most marginalised households to enable them to pay for school fees and purchase school items.

While it is important to bear in mind the limitations associated with poor quality attendance data, qualitative endline evidence from project evaluation reports suggests that **interventions facilitating access to loans and savings and supporting income-generation are helping to keep girls in school.** Village savings and loan associations, such as those implemented by IRC (DRC) and WV (Zimbabwe), were found to increase the financial capacity of

---

[74] Shonchoy, Abu & Kudo, Yuya & Takahashi, Kazushi. (2017). Can Solar Lanterns Improve Youth Academic Performance? Experimental Evidence from Bangladesh. The World Bank Economic Review.10.1093/wber/lhw073; Furukawa, Chishio (2014). Do Solar Lamps Help Children Study? Contrary Evidence from a Pilot Study in Uganda. The Journal of Development Studies, 50(2):319-341.

caregivers and their ability to pay school fees in a timely manner, through training support members to set up their own businesses, or budgeting support.

Evidence from IRC (DRC) and Child Hope (Ethiopia) also suggests these **interventions, including livelihood programmes, are encouraging caregivers to spend more on girls' education.** At endline, IRC reported that 81 per cent of EA$E members (the savings and loan association they set up) declared that they spend more on their children's education. However, this could also reflect caregivers spending more as a result of increased costs associated with schooling as girls get older. In Child Hope (Ethiopia) project areas, a similar pattern emerged, which outdid its endline output-level target in the proportion of caregivers indicating an increase in their expenditure on girls' education (81 per cent). Qualitative data in their endline report found that, as at baseline and midline, the livelihoods programme empowered mothers to advocate for their daughters' education, which may explain the positive and significant association found (at the 10 per cent level) between girls whose primary caregivers are in receipt of livelihood support and their school attendance.

One project (CfBT (Kenya)) provided cash transfers to marginalised households. Qualitative evidence from their report points to households' use of cash transfers to not only meet the costs of school expenses, but also to finance basic needs, such as food and electricity, which is helping to stave off hunger and improve girls' ability to complete homework in the evenings. This is corroborated by quantitative data which reveals that the mean attendance of cash transfer beneficiaries increased from 88 per cent at midline to 94 per cent at endline. This compares with a more modest increase in mean attendance for girls in intervention schools more generally, which increased from 84 per cent at midline to 85 per cent at endline. Research reviewed by 3ie in Snilstveit *et. al* (2016: 21)[75] finds that cash transfers have substantive effects on the ability to participate in school, with positive effects on school enrolment, attendance, completion and drop out.

There was some concern, however, that these **interventions are not always able to reach their target beneficiaries, with marginalised families and communities unlikely to benefit.** In IRC (DRC), qualitative findings revealed that the popularity of the loans led to a saturated market, which in turn led to some difficulties for members in getting their loans reimbursed and discouraged others from participating in the scheme. In STC (Ethiopia), evidence presented at endline highlighted that VSLA groups were not successful in targeting the poorest members, since they often could not save a sufficient level of money to qualify and be part of the group. The limited and short-term nature of improvements from microfinance, as well as its often exclusionary nature, have been highlighted by other studies and reviews such as Banerjee et al, 2015 and Duvendack and Mader (2017).[76]

School-related factors

All SCW projects address school-related factors within their theory of change as key barriers to girls' education. Projects at endline reported improvements in school facilities and access (eight projects), provision of teachers and teaching materials (seven projects), quality of teaching (six projects), and school environment (five projects). One project reported an improvement in school governance (Camfed (Zimbabwe-Tanzania)), although there was no evidence of any direct impact on learning outcomes.

Across the window, we find wide reference to improvements in school facilities, sanitation and in the availability of school materials, with the provision of learning materials in some instances having positive effects on literacy and numeracy outcomes (three projects), and infrastructure and sanitation facilities improvements helping to encourage more regular attendance of girls at school (three projects). Two of the three projects providing accelerated learning and alternative schooling presented evidence linking these to improvements in literacy and numeracy.

---

[75]Snilstveit, B, Stevenson, J, Menon, R, Phillips, D, Gallagher, E, Geleen, M, Jobse, H, Schmidt, T and Jimenez, E, 2016. The impact of education programmes on learning and school participation in low- and middle-income countries: a systematic review summary report, 3ie Systematic Review Summary 7. London: International Initiative for Impact Evaluation (3ie)

[76] Banerjee, Abhijit, Esther Duflo, Rachel Glennerster, and Cynthia Kinnan. 2015. "The Miracle of Microfinance? Evidence from a Randomized Evaluation." American Economic Journal: Applied Economics, 7(1): 22-53; Duvendack, M., Mader, P. (2017) Poverty Reduction or the Financialisation of Poverty?, in Seduced and Betrayed: Exposing the Contemporary Microfinance Phenomenon. University of New Mexico Press ISBN 978-0-8263-5796-0 (Chapter (peer-reviewed)

Key findings

Classroom infrastructure and sanitation facilities

- There have been widespread reported improvements in classroom infrastructure and in the quality of learning materials, although there are persistent concerns about the adequacy of sanitation facilities, including access to water.

- There is some evidence that infrastructure and facilities improvements are helping to encourage more regular attendance.

Long distances

- The long distances to schools remains a barrier to girls' ability to attend school.

Educational resources

- Learning materials are promoting independent learning, with positive effects on student motivation and learning outcomes.

- Nevertheless, learning resources risk being overly relied upon as a substitute for teaching.

- The availability and supply of learning resources remains insufficient, and threatens to undermine enrolment outcomes, particularly for marginalised households.

- School materials may not always be reaching the intended beneficiaries.

Teaching quality and methodologies

- Although projects found improvements in the use of participatory teaching practices and pedagogies since midline, teachers could not always recall the specific training they had received.

- Improved teaching methods and methodologies are helping to engage girls in class, however only one project demonstrated the link between these and improvements in learning outcomes.

Learning environment and accelerated learning schools

- There has been some increase in the psychological support offered to girls at school, but the effects of these on educational outcomes are inconclusive.

- Accelerated learning and alternative schools are helping girls catch up with learning.

## School facilities and sanitation

Across the window there were widespread reported improvements in classroom infrastructure. In Child Hope's (Ethiopia) project, which constructed libraries, reading corners, additional classrooms and girls' toilets, caregivers were significantly more satisfied with classrooms at endline than at midline, with greater improvements for the treatment group. There was also a significant increase since midline in the proportion of caregivers observing that classrooms had improved over the past two years, with no increase found in the control group. BRAC (Afghanistan), AKF (Afghanistan), ACTED (Afghanistan) and WUSC (Kenya) all presented evidence in their endline evaluation reports of improvements and increased satisfaction with classroom infrastructure. However, none of the projects presented quantitative evidence about the impacts of these infrastructure interventions on attendance or learning outcomes.

Persistent overcrowding may hamper improvements in educational outcomes. In their endline evaluation report, ACTED (Afghanistan) found over 20 per cent of respondents reporting that classrooms were 'more crowded', despite over 90 per cent witnessing an increase in the overall number of classrooms. ACTED's project design and targets, as well as their infrastructure plans since, did not assume a demand for education that would surpass targets by over double the amount of initial target beneficiaries. Qualitative evidence found that, despite investments in infrastructure, there was an increase in demand for girls' education during the course of the project, which resulted in classroom overcrowding. Similar observations were noted in Child Hope's (Ethiopia) and WUSC's (Kenya) projects, suggesting that learning outcomes in the long-term could be impeded if the development of

facilities are not able to keep up with increases in demand for education, although other studies (e.g. Jones, 2016[77]) suggest that there is little effect from overcrowding alone.

Box 8: Classroom quality – what has changed overall?

Findings presented below are based on primary data gathered by the EM at endline. Treatment averages are shown on top and control averages are shown on the bottom. Statistically significant changes are indicated in the text below.

Crowdedness of girl's class

— Treatment ⋯⋯ Control
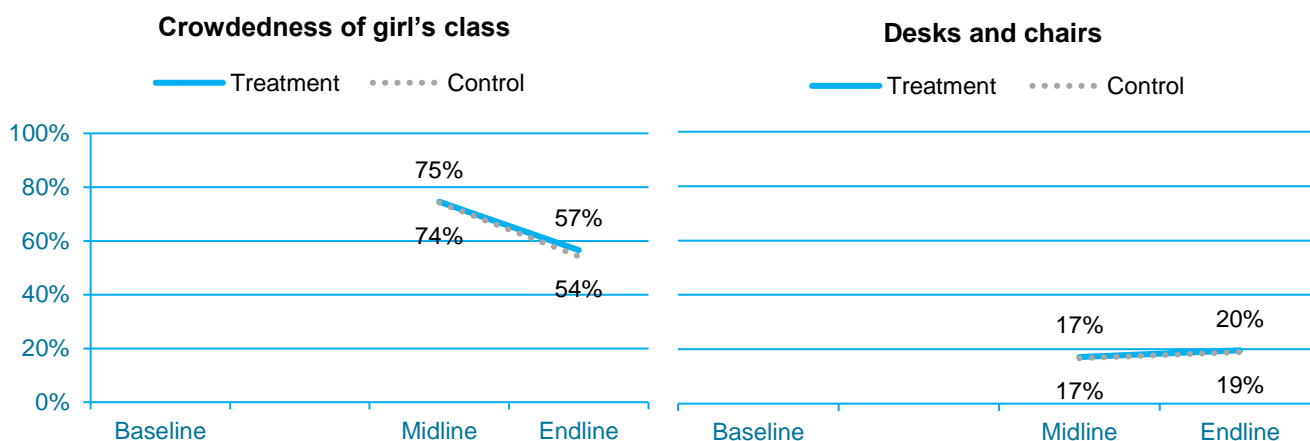
School Visit Survey (Evaluation Manager) | "There are 30 children or more in girl's class". Average across all SCW EM project areas, excluding Camfed (Zimbabwe – Tanzania), Afghanistan and Sierra Leone areas.

Desks and chairs

— Treatment ⋯⋯ Control

Household Survey (Evaluation Manager) | "More/better desks or chairs". Proportion of caregivers that answered "yes". Average across all SCW EM project areas, excluding Camfed (Zimbabwe – Tanzania).

Data gathered by the EM showed that, in contrast with findings from some project reports, **girls' classrooms have become less crowded over time.** At endline, fewer treatment and control schools were found to have 30 children or more in the girl's class than at midline, although changes were not statistically significant (75% / 57%, 74% / 54%, -2 pp).

This could be attributed to **improvements in the availability and quality of classroom infrastructure,** which resonates with findings from projects' endline reports. At endline, a high proportion of caregivers indicated that there were "more/better desks or chairs" than at midline, across both treatment and control areas (17% / 20%, 17% / 19%, 1 pp).

Classroom quality – evidence from projects' data

Findings presented below are based on primary data gathered by projects at endline. Projects presented below are those that at a minimum collected data for treatment areas from midline to endline.

---

[77] Sam Jones, How does classroom composition affect learning outcomes in Ugandan primary schools?, In International Journal of Educational Development, Volume 48, 2016, Pages 66-78.
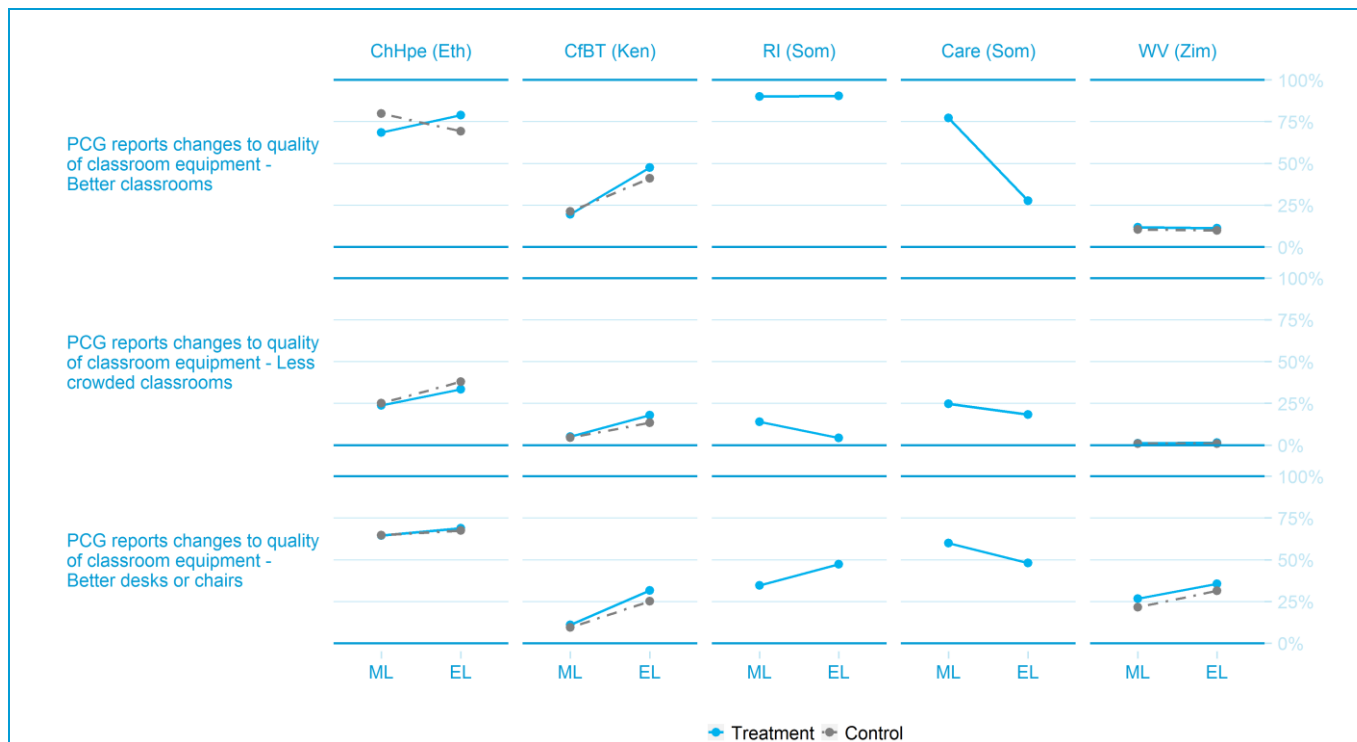
At endline, caregivers' perceptions of the quality of classrooms were mixed. While Child Hope (Ethiopia) and CfBT (Kenya) reported an increase in the number of caregivers indicating better quality classrooms and furniture, this was not the case in CARE (Somalia), where fewer caregivers reported classroom improvements than at midline.
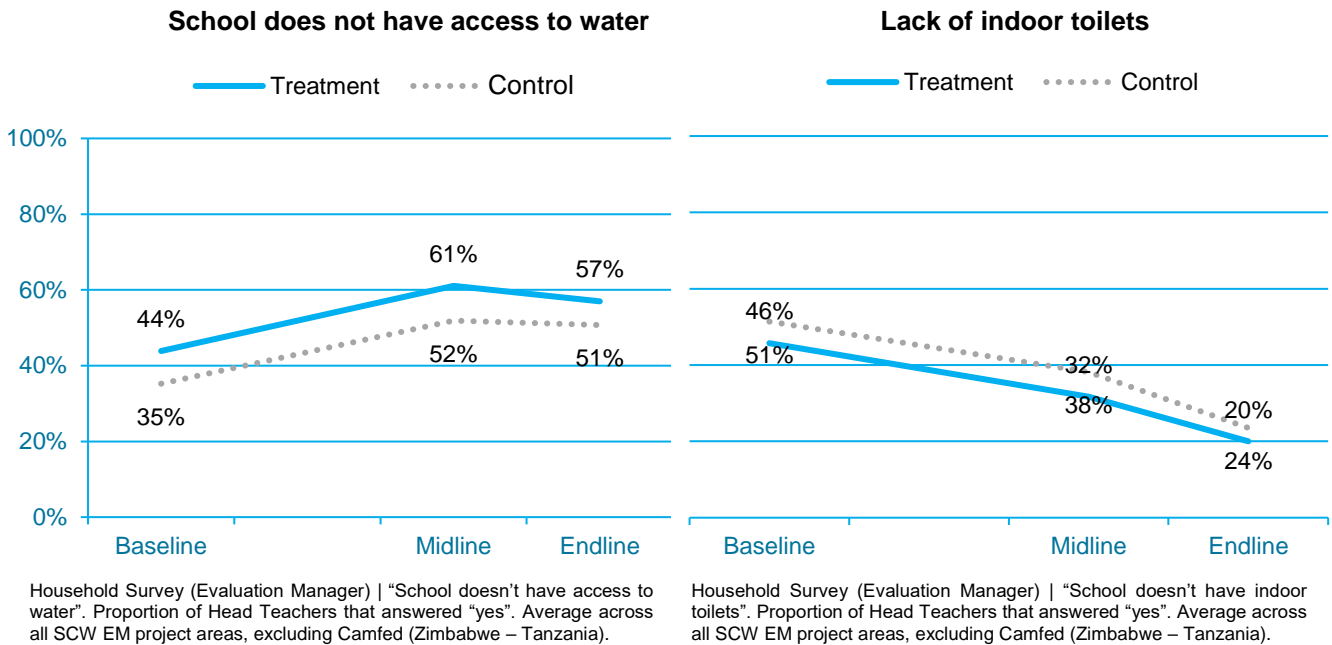
In RI (Somalia), there was little to no change in perceptions of classroom quality, although this was already high at midline. Although Child Hope (Ethiopia) cautioned against classroom overcrowding in their endline report, primary data from the project suggested that fewer caregivers perceived classrooms to be crowded than at midline.

Access to adequate sanitation and WASH facilities remains a persistent threat to girls' attendance. Despite improvements in the availability of segregated toilets, BRAC (Afghanistan), AKF (Afghanistan), Child Hope (Ethiopia) and WUSC (Kenya) all revealed continued dissatisfaction among girls with the quality of sanitation facilities. There was some evidence to suggest that this was because of a **scarcity of clean water**, which projects failed to address when constructing toilets and latrines. Qualitative evidence presented by Child Hope found evidence of students begging for water near school during their breaks and unable to use the toilet constructed by the project. Water shortages could in the long term deter girls from school since adequate sanitary facilities were found by the project to be particularly important for girls during menstruation.

Box 9: Sanitation facilities – what has changed overall?

Findings presented below are based on primary data gathered by the EM at endline. Treatment averages are shown on top and control averages are shown on the bottom. Statistically significant changes are indicated in the text below.

In line with findings harvested from project reports, data gathered by the EM also revealed important differences in trends around **access to water and the availability of toilets** at the window-level. While there has been a steady decrease in the proportion of schools without access to indoor toilets from baseline to endline, access to water has worsened over time for both treatment and control areas. This suggests that while projects may have been effective in providing indoor toilet facilities, they have neglected to address basic infrastructure needs, which themselves may be hampered by contextual factors (such as drought).

School does not have access to water

——— Treatment   ······ Control

Lack of indoor toilets

——— Treatment   ······ Control

Household Survey (Evaluation Manager) | "School doesn't have access to water". Proportion of Head Teachers that answered "yes". Average across all SCW EM project areas, excluding Camfed (Zimbabwe – Tanzania).

Household Survey (Evaluation Manager) | "School doesn't have indoor toilets". Proportion of Head Teachers that answered "yes". Average across all SCW EM project areas, excluding Camfed (Zimbabwe – Tanzania).

This confirms findings reported by the EM at both baseline and midline, which concluded that issues such as poor access to water and other basic conditions for education and learning need to be addressed in order to enable projects to deliver improvements successfully[78].

There is some evidence linking the provision of girls' latrines at school to improved attendance outcomes.
RI (Somalia) found that the presence of girls' latrines at school had a positive significant effect on attendance[79]. Qualitative evidence cited reduced shyness among girls, which may be a factor in helping girls feel more comfortable at school and therefore in attending more regularly:

> *Girls were not getting toilets in the previous time, and the school used to have two public toilets. Now the organization built separate toilets for the girls, and that has eliminated their shyness, and they are satisfied for learning.*

> Government official, Puntland
> Endline evaluation report, RI (Somalia)

This is confirmed by quantitative evidence, with the proportion of girls who feel comfortable using latrines increasing from 53 per cent at midline to 61 per cent at endline. CARE (Somalia), which also constructed latrines, also found that menstruation was mentioned less frequently in interviews and focus group discussions as a barrier to attendance at endline, despite the increase in the proportion of girls of menstruating age from midline to endline (from 39 per cent to 51 per cent). Qualitative interviews and group discussions from Child Hope (Ethiopia) were similarly reported to reference girls' toilets as helping to improve attendance, since it made it easier for them to attend school during menstruation.

---

[78] GEC SCW Midline Evaluation Report, 2017: 66, 100
[79] According to their endline evaluation report, the presence of girls' latrines at a school had a positive significant correlation with attendance.

Long distance to school

Long distance to school was cited as a barrier to girls' attendance and learning outcomes in school by a number of projects, both because walking long distances to school discourages students from attending regularly or concentrating in class, and because of the potential safety and harassment issues associated with unaccompanied girls travelling to school. This finding is supported by experimental studies such as Burde and Linden (2013) which suggest that girls are particularly sensitive to distance[80].

While projects acknowledged the long distance to schools as a barrier, very few projects explicitly addressed this in their theory of change. Three projects aimed to address this by providing classes or education facilities in girls' communities close to their homes (BRAC (Afghanistan), ACTED (Afghanistan) and STC (Ethiopia).

Encouraging evidence from WV's (Zimbabwe) evaluation report suggests that at endline, caregivers and girls in the treatment group who received a bicycle through its Bicycle Education Empowerment Programme (BEEPs) were significantly more likely to be enrolled in school at endline and report that the bikes improved their school attendance more than in the control group (see also Muralidharan and Prakash (2017)[81] who highlight the benefits of this low cost intervention in relation to enrolment, attendance and learning outcomes among female secondary school pupils in Bihar).

While there is limited quantitative data linking this to learning outcomes, qualitative evidence from their report suggests that reducing the tiredness of girls may help improve attendance and learning, while also benefiting other children being given lifts.

There is little evidence of the effect of interventions addressing distance to school on girls' attendance or learning outcomes. While there was strong evidence that the distance barrier had been addressed in BRAC's (Afghanistan) project areas – as evidenced both by mothers and girls reporting the ease of accessing school, as well as shorter overall journey times to school noted by the project, in ACTED's (Afghanistan) evaluation report, qualitative findings suggested that *distance to facilities* remained a main barrier to attendance, with over 50 per cent of respondents citing the distance travelled as a reason for what makes the journey to school unsafe. However, ACTED found no direct relationship between distance to school and the ASER scores of their targeted girls.

Lack of school materials

Beyond school infrastructure, there are a number of materials and learning resources associated with schooling, such as textbooks, libraries and classroom supplies, which can encourage positive learning and attendance outcomes. Most projects identified a lack of these school materials within their ToCs, and sought to provide them directly during the course of their intervention.

At endline, we find that there have been widespread reported improvements in the quality of learning materials. AKF (Afghanistan), ACTED (Afghanistan), WUSC (Kenya), CARE (Somalia) and Plan (Sierra Leone) all found improvements with the quality of learning materials, with some reporting increased satisfaction with the availability or quality of the resources.

Among items provided as part of the learning resources that were supplied, there was some evidence that **interventions promoting independent learning had positive effects on learning outcomes.** Two projects (WV (Zimbabwe) and Camfed (Zimbabwe and Tanzania)) presented evidence of improvements in literacy and numeracy linked to the provision of learning resources (see further below), with ACTED (Afghanistan) presenting

---

[80] Burde, Dana, and Leigh L. Linden. 2013. "Bringing Education to Afghan Girls: A Randomized Controlled Trial of Village-Based Schools." American Economic Journal: Applied Economics 5(3): 27-40. For example, the enrolment rate for girls in the study fell by 19 percentage points per additional mile, while boys' enrolment only fell by 13 percentage points. Similarly the study finds that girls' test scores fell by 0.24 standard deviations per additional mile while boys fell by 0.15 standard deviations.

[81] Karthik Muralidharan, Nishith Prakash (2017). Cycling to School: Increasing Secondary School Enrollment for Girls in India. American Economic Journal: Applied Economics 9(3): 321 – 50.

qualitative evidence supporting this. Research from Snilstveit *et al.* (2016: 29)[82] supports these findings, suggesting that the availability of learning materials can help children engage with the curriculum and promote self-study.

Camfed (Zimbabwe and Tanzania) provided study guides in Maths, English, Science (and in *How to Learn English* in Tanzania). Between midline and endline the report noted that study guides were associated with higher English attainment scores of between five and six percentage points in both Tanzania and Zimbabwe. The guides were also associated with higher maths attainment scores, although to a lesser extent in Zimbabwe and a slightly greater extent in Tanzania.

WV (Zimbabwe) provided books and accompanying training in how to use the books through the *Happy Readers* programme, which aimed to improve the quality of literacy and numeracy instruction that girls receive. Although no significant effects were noted in literacy for the treatment group overall since midline, there were significant effects in particular districts on certain tests (such as the letters and sounds (EGRA1) and comprehension (EGRA5) tests in Mangwe), which qualitative evidence attributes to the *Happy Readers* programme.

Qualitative evidence from across these projects found that this was because **the provision of books supported independent learning and increased student motivation to learn, which in turn had positive effects on learning outcomes.** In both Camfed (Zimbabwe and Tanzania) and WV (Zimbabwe), qualitative findings attributed these improvements to being able to use the books at home, which improved studying habits and increased motivation to learn when in class.

In contrast, there is some evidence that learning resources risk being overly relied upon as a substitute for teaching. In RI (Somalia), mobile libraries were significantly associated with lower learning scores among girls at school that received mobile libraries, which the report attributes to the fact that these schools may have been generally underperforming compared to schools that did not get mobile libraries. While this may suggest that mobile libraries were being effectively targeted in RI project areas, a similar finding was observed in Camfed's (Zimbabwe and Tanzania) project, where daily use of the study guides were associated with literacy scores around two percentage points lower than for other users. At midline, Camfed found that such daily use was associated with less well-resourced schools for whom the study guides provided an invaluable teaching resource. This suggests that while learning resources such as study guides and books can support independent learning, they risk hampering learning outcomes if used as a substitute for good quality teaching and instruction.

Box 10: Teaching methods and resources for teaching – what has changed overall?

Findings presented below are based on primary data gathered by the EM at endline. Treatment averages are shown on top and control averages are shown on the bottom. Statistically significant changes are indicated in the text below.

---
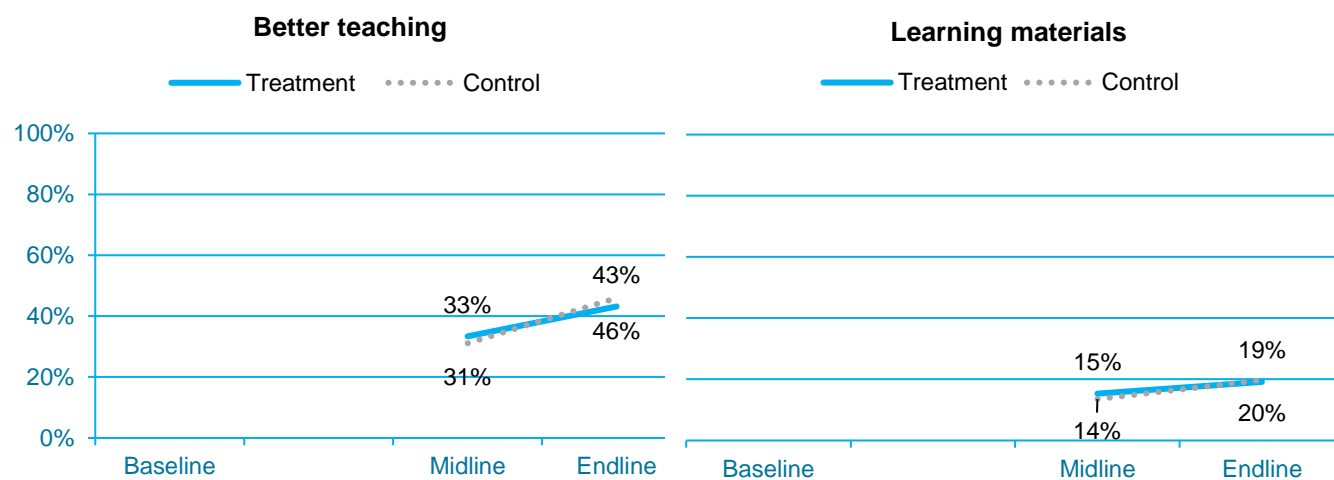
[82] Snilstveit, B, Stevenson, J, Menon, R, Phillips, D, Gallagher, E, Geleen, M, Jobse, H, Schmidt, T and Jimenez, E, 2016. The impact of education programmes on learning and school participation in low- and middle-income countries: a systematic review summary report, 3ie Systematic Review Summary 7. London: International Initiative for Impact Evaluation (3ie)

Household Survey (Evaluation Manager) | "Better teaching". Proportion of caregivers that answered "yes". Average across all SCW EM project areas, excluding Camfed (Zimbabwe and Tanzania).

Household Survey (Evaluation Manager) | "Better/more learning materials". Proportion of caregivers that answered "yes". Average across all SCW EM project areas, excluding Camfed (Zimbabwe and Tanzania).

Endline primary data gathered by the EM chimes with findings from project endline reports that suggest **improved teaching and learning methods are helping to engage girl students**. Although no significant effects were observed on the quality of teaching or learning materials, the share of caregivers stating that schools now have "better teaching" has increased since midline for both treatment and control areas, with a larger effect observed among control areas (33% / 43%, 31% / 46%, -5 pp). In terms of learning materials, the share of caregivers stating that there are now "better/more learning materials" has more or less remained stable for treatment and control groups (15% / 19%, 14% / 20%, -2 pp).

Teaching methods and resources for teaching – evidence from projects' data

Findings presented below are based on primary data gathered by projects at endline. Projects presented below are those that at a minimum collected data for treatment areas from midline to endline (except for IRC (DRC), which did not have data available for "new teaching methods", but did for "teaching quality").

In IRC's (DRC), Child Hope's (Ethiopia) and RI's (Somalia) projects, perceptions of teaching quality among treatment areas were already very high at midline, with little to no additional changes reported at endline (see graph below). For CfBT's (Kenya) project, where teachers received coaching on literacy and gender-friendly pedagogy, there was an increase in the proportion of caregivers perceiving improvements in the use of teaching methods, although no differences were found between control and treatment areas.
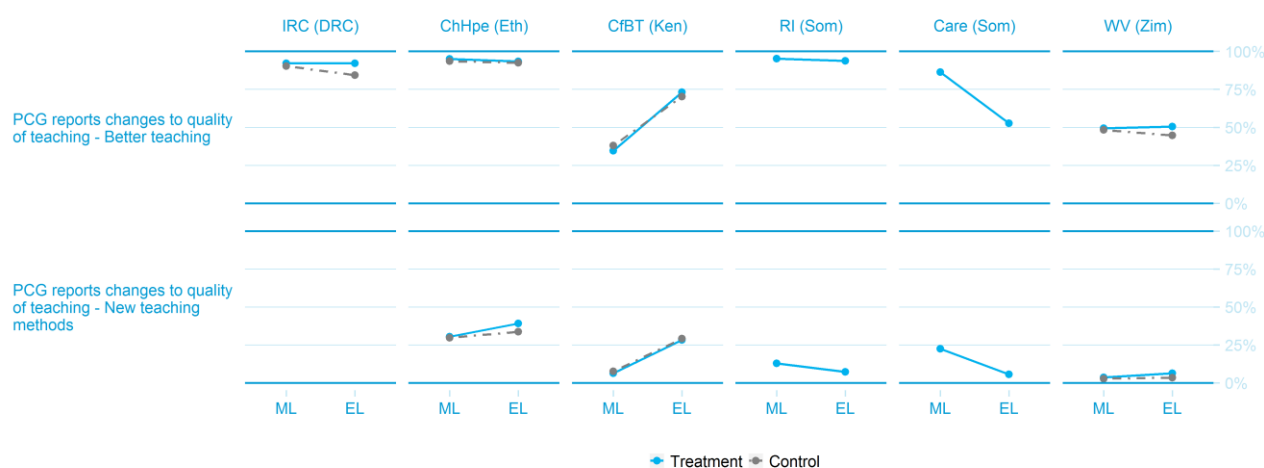
Nevertheless, the availability and supply of learning resources remains insufficient, with some evidence that this may undermine enrolment outcomes, particularly for girls from marginalised households. Both AKF (Afghanistan) and ACTED (Afghanistan) identified the inability of households to pay for school materials as a demand-side barrier, putting girls' enrolment at risk for those households unable to afford them. AKF's parents stated that having to buy such resources was a significant burden for poor families, with the provision of stationery pivotal for households in deciding whether or not to send their girls to school. WUSC (Kenya) and CARE (Somalia) both also spoke of the persistent limited availability of textbooks, with reports of students continuing to share texts while in class.

There is also some evidence that school materials may not always be reaching the intended beneficiaries. For example, qualitative evidence from WV (Zimbabwe) found that teachers did not always distribute the *Happy Readers* books during lessons. Wider research supports these findings, with evidence from other programmes that textbooks have sometimes been kept in storage and not distributed to students (Snilstveit *et al*, 2016: 30)[83].

---

[83] Snilstveit, B, Stevenson, J, Menon, R, Phillips, D, Gallagher, E, Geleen, M, Jobse, H, Schmidt, T and Jimenez, E, 2016. The impact of education programmes on learning and school participation in low- and middle-income countries: a systematic review summary report, 3ie Systematic Review Summary 7. London: International Initiative for Impact Evaluation (3ie)

Isolated observations from girls cited in ACTED's endline evaluation report reveals that that some of the stationery supplies and textbooks provided by the program were being seized by school management, and therefore did not always reach the girls themselves.

Teachers' inadequate pedagogy and teaching skills

Most projects identified poor quality teaching as a key barrier to girls' educational outcomes, including teachers' poor knowledge or grasp of English and maths and ineffective classroom management or teaching skills, which can result in girls' educational needs being overlooked. In response, projects aimed to address these gaps through various teacher training modules, focusing on basic literacy and numeracy training; gender-sensitive and gender-responsive pedagogies; and participatory or child-centred learning techniques.

A number of projects found improvements in the use of participatory teaching practices and pedagogies since midline. BRAC (Afghanistan), which trained female teachers in the use of participatory learning techniques such as quizzes, role play and demonstration, observed an increase in the proportion of CBGS teachers using the new classroom methodologies from 55 per cent at midline to 75 per cent at endline, with a similar increase for government school teachers (from 61 per cent at midline to 81 per cent at endline). AKF (Afghanistan), ACTED (Afghanistan), IRC (DRC) and Child Hope (Ethiopia) all observed similar improvements in the use of new classroom methodologies.

Although classroom observations revealed an increase in the prevalence of participatory teaching practices across projects, **teachers could not always recall the specific training they had received.** Both ACTED and BRAC (Afghanistan), which provided training to teachers in new teaching methods, found in their qualitative research that most teachers failed to identify the nature of the training undertaken, or how they had implemented the methods they had learnt, with some reporting to not have received any training through the projects.

There is some evidence to suggest these methods and methodologies are helping to engage girl students in class. WUSC (Kenya), for example, which provided teacher training in girl-friendly approaches and methods, found from speaking with teachers and girls that there was increased participation and confidence among girls to answer questions and seek help in class where needed, with more than three quarters (80 per cent) of in-school girls reporting some improvement in the way teachers interact and engage with students.

Similarly, in CfBT's (Kenya) project areas, where teachers received coaching on literacy and gender-friendly pedagogy, qualitative evidence from girls suggests their engagement may be due to increased motivation and self-esteem:

> *Some teachers are very good in class. They praise us and show us how to behave well. Boys are advised to respect girls…*

> Girl student, County 62
> Endline Evaluation Report, CfBT (Kenya)

ACTED (Afghanistan) and Child Hope (Ethiopia) also reported similar observations from teachers in their endline reports, with evidence of girls being more engaged and active in class as a result of participatory methods.

Only one project implementing teacher training and support interventions demonstrated conclusive effects on girls' education. Most projects implementing these kinds of interventions did not sufficiently or reliably demonstrate the links between these interventions and their effects on learning outcomes. However, Child Hope (Ethiopia), which trained teachers in active-learning methods to improve teaching in literacy and numeracy, found that half of girls (51 per cent) who attended teacher tutorials improved their literacy since midline (any increase in literacy scores between midline and endline), with three out of five girls (63 per cent) improving their numeracy (any increase in numeracy scores between midline and endline). Qualitative reports from girls suggests that this was because of the small classes and the direct support given to them during the tutorials to help them grasp difficult concepts. Although ACTED (Afghanistan) identified a link between higher numeracy and literacy scores and improved teacher practices, there are reservations about the quality of evidence as this was only presented anecdotally.

School environment

Some projects identified the lack of psychological support to girls at school as a barrier to their education, with girls whose needs are not met at risk of dropping out or achieving lower educational attainment.

At endline, **projects presented some evidence of an increase in the psychological support offered to girls at school.** STC (Ethiopia), for example, reported that the proportion of girls indicating that their psychological needs were met increased from 29 per cent at midline to 83 per cent at endline. Similarly, Child Hope (Ethiopia) reported that at endline, 71 per cent of targeted marginalised girls who had received psycho-social support from the counsellors reported that it helped them cope with their problems.

There was limited evidence of changes in educational outcomes as a result of these interventions, however. WUSC (Kenya), which set up Counselling Clubs and adopted a 'cascade' model of hiring trained Counsellors to train teachers and parents in counselling, found that 58 per cent of girls reported someone other than a family member talking to them about enrolling or staying in school over the past two years. However, none of the projects presented conclusive evidence linking these interventions to a change in girls' learning or attendance outcomes.

Accelerated learning and alternative schools

Accelerated learning and alternative schools are proving effective in improving girls' learning and attendance. AKF (Afghanistan) and IRC (DRC) both established accelerated learning programmes (ALPs) to enable girls to progress through an accelerated cycle of primary schooling. Both projects presented strong evidence at endline that these programmes helped increase girls' enrolment and learning outcomes.

AKF (Afghanistan) established ALP classes as part of the wider community-based education it delivered in areas without access to government schools. At endline, the project found higher literacy and numeracy scores among re-contacted ALP girls as compared to their government school counterparts in the same grade, with improvements in numeracy outcomes as compared with midline. Girls in ALP classes also have higher attendance at endline. These improvements may be attributed to the quality of teaching, since classroom observations found that teachers in community-based classes (including ALPs) displayed better use of learner-centred and gender-fair methodologies than teachers in government school classes, alongside more respectful behaviours towards students.

In IRC (DRC), ALP classes were available at no cost to out-of-school girls and boys to enable them to complete a full cycle of primary school. Qualitative data found that ALP classes helped mitigate the economic barriers to education:

> *I studied at the primary school of Moleka but was sent home since I lacked money [to pay fees]. When I went home I met the director of the [ALP] center, who spoke to me about the center, and I informed my parents who enrolled me here.*
>
> Girl student, Equateur
> Endline evaluation report, IRC (DRC)

At endline the report finds that girls attending these classes from midline to endline had higher literacy and numeracy scores as compared to girls who were out of school, with additional positive effects observed in enrolment and drop-outs.

Attitudes towards girls' education

Around half of SCW projects addressing negative attitudes towards girls reported improvements in these barriers, including in: attitudes towards girls' education (two projects)**,** awareness of the value of education (three projects), and increased family (seven projects) and community (five projects) support for girls and their education.

However, project data often presented evidence of attitudinal shifts in general terms, with their reports addressing a number of related sub-barriers at any one time. As such, we report on findings at two levels, looking at evidence in support of changed attitudes among households (household-level) and community members (community-level).

Across the window, we find in particular that community and parental attitudes are improving, with some evidence of community-based interventions having positive effects on attendance (six projects).

Key findings

Community and parental attitudes

- At endline there are reported improvements in parental and community attitudes towards girls' education, with some evidence of positive effects on attendance.

- Mothers are acting as 'troubleshooters' and helping to mobilise resources to identify and overcome barriers to girls' enrolment in school.

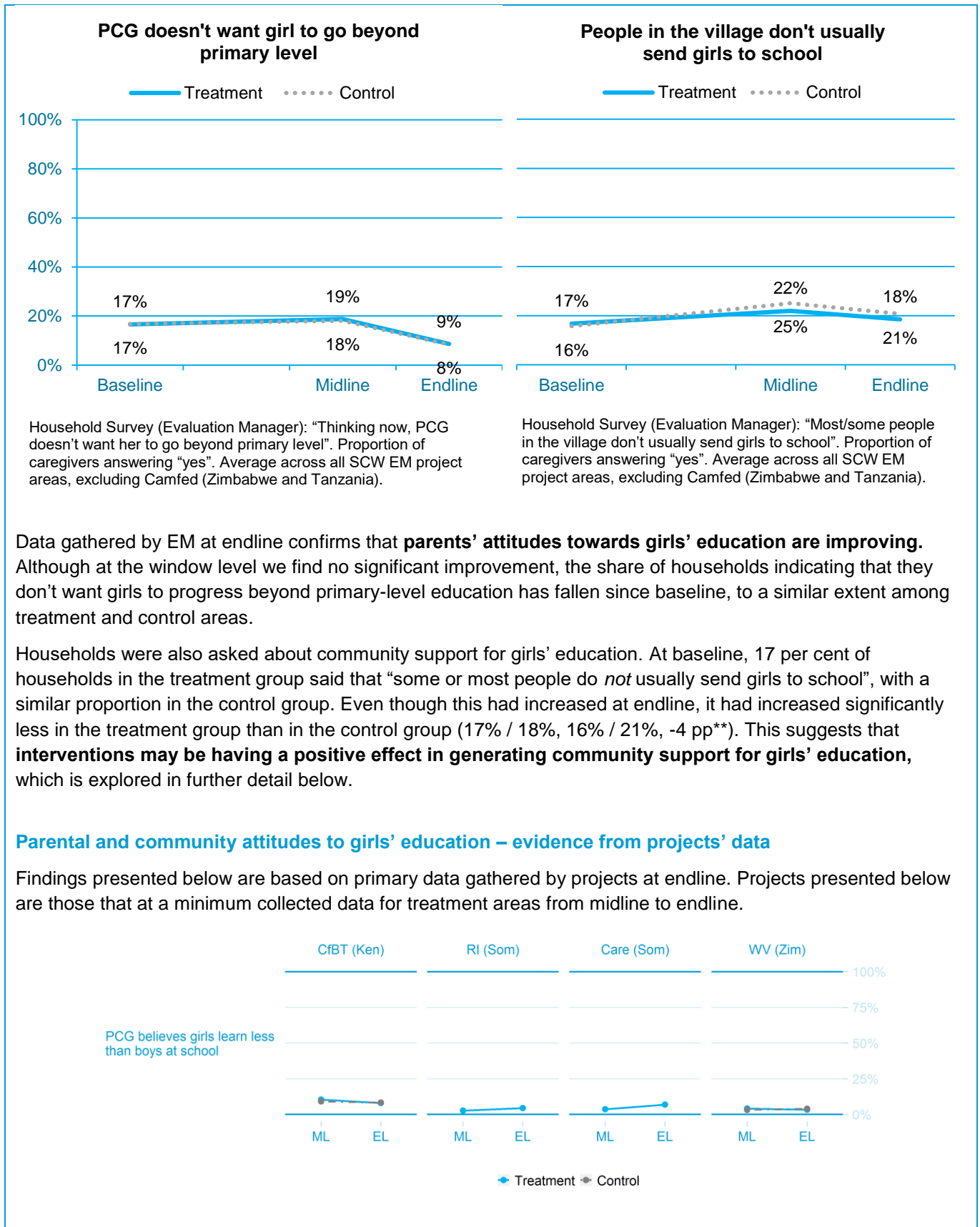*Community and caregiver support for girls' education*

- Adult literacy classes are helping to improve caregivers' perceptions of the value of girls' education.

- Community workshops are pivotal in dispelling myths related to girls' education, while community members and mothers are playing an active role in monitoring girls' attendance.

- Radio messaging on girls' education can be effective in improving enrolment and attendance, providing access barriers are addressed.
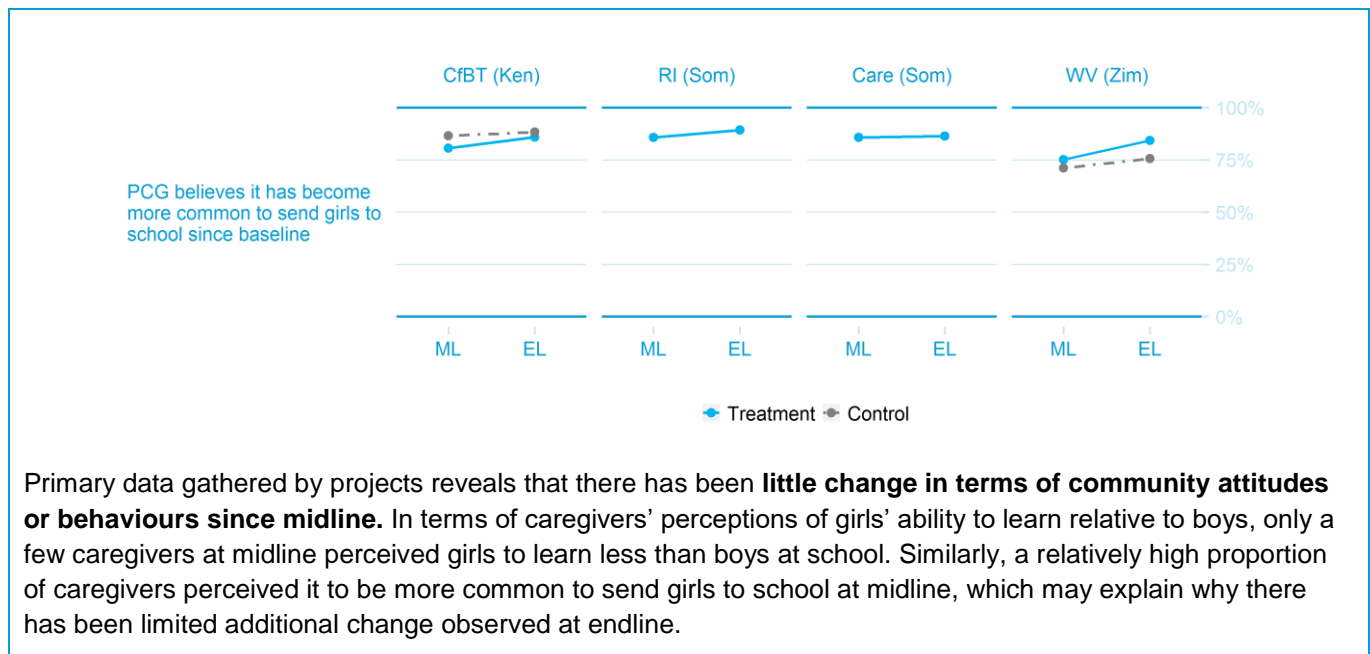

Household-level barriers

At endline there was widespread reports of improved parental attitudes towards girls' education. Media messaging, Mothers' Groups, and community meetings are all helping to promote positive beliefs about the importance of girls receiving an education. BRAC (Afghanistan), AKF (Afghanistan), Camfed (Zimbabwe and Tanzania), Save the Children (Ethiopia), RI (Somalia), CARE (Somalia) and Child Hope (Ethiopia) all presented quantitative evidence of improved perceptions among parents and caregivers towards girls' education.

Box 11: Parental and community attitudes to girls' education – what has changed overall?

Findings presented below are based on primary data gathered by the EM at endline. Treatment averages are shown on top and control averages are shown on the bottom. Statistically significant changes are indicated in the text below.

PCG doesn't want girl to go beyond primary level

Household Survey (Evaluation Manager): "Thinking now, PCG doesn't want her to go beyond primary level". Proportion of caregivers answering "yes". Average across all SCW EM project areas, excluding Camfed (Zimbabwe and Tanzania).

People in the village don't usually send girls to school

Household Survey (Evaluation Manager): "Most/some people in the village don't usually send girls to school". Proportion of caregivers answering "yes". Average across all SCW EM project areas, excluding Camfed (Zimbabwe and Tanzania).

Data gathered by EM at endline confirms that **parents' attitudes towards girls' education are improving.** Although at the window level we find no significant improvement, the share of households indicating that they don't want girls to progress beyond primary-level education has fallen since baseline, to a similar extent among treatment and control areas.

Households were also asked about community support for girls' education. At baseline, 17 per cent of households in the treatment group said that "some or most people do *not* usually send girls to school", with a similar proportion in the control group. Even though this had increased at endline, it had increased significantly less in the treatment group than in the control group (17% / 18%, 16% / 21%, -4 pp**). This suggests that **interventions may be having a positive effect in generating community support for girls' education,** which is explored in further detail below.

Parental and community attitudes to girls' education – evidence from projects' data

Findings presented below are based on primary data gathered by projects at endline. Projects presented below are those that at a minimum collected data for treatment areas from midline to endline.

Primary data gathered by projects reveals that there has been **little change in terms of community attitudes or behaviours since midline.** In terms of caregivers' perceptions of girls' ability to learn relative to boys, only a few caregivers at midline perceived girls to learn less than boys at school. Similarly, a relatively high proportion of caregivers perceived it to be more common to send girls to school at midline, which may explain why there has been limited additional change observed at endline.

Mothers are effectively mobilising support for girls' education and helping to enrol out-of-school girls.
Endline reports from BRAC (Afghanistan), Save the Children (Mozambique) and WV (Zimbabwe) all found evidence of mothers working with other parents and caregivers to raise awareness of girls' education.

BRAC (Afghanistan) organised Mothers' Forums for mothers to discuss issues related to girls' education and ensure their regular attendance. Evidence presented at endline found that mothers felt empowered to mobilise other mothers in the community who did not have daughters enrolled in school, and were effective in raising awareness of benefits of education that they had discussed in the Mothers' Forum. In addition, qualitative evidence from male community members and CBGS girls pointed to the ability of Mothers' Forums to help overcome individual barriers to girls' education and ensure the girls attend regularly.

Qualitative data from WV (Zimbabwe) found that mothers within their Mothers' Groups acted as 'troubleshooters', following up on girls absent from school and mobilising resources and other stakeholders to identify and help overcome barriers to enrolling their girls in school. As one mother noted:

> Through IGATE a group known as the Mothers Group was formed and they are there to encourage the girl child to be in school […] As a member of the Mothers Group, it is my responsibility to check on girls who are not going to school in the community […] I have to approach the parents and guardians of that particular girl in order to find out the reasons why. […] I will then encourage them to come through and talk to the Mothers Group as well as the school authorities. We are normally available at school on Thursdays so that they can present their case to the school before the child is struck off the school register […]. As members of the Mothers Group we find ways to ensure that the child gets pens and books to use in school.

Caregiver, Gokwe South
Endline Evaluation Report, WV (Zimbabwe)

According to their endline report, caregivers involved in the groups were significantly more likely to have a child enrolled in school at endline, while girls who reported that their family member was involved in MG activities were significantly more likely to score higher in literacy.

There were nevertheless some concerns that Mothers' Groups may not be as effective in addressing wider issues related to girls' education, such as early marriage or financial barriers. Mothers in BRAC (Afghanistan) reported feeling powerless in wider barriers such as household poverty or the diminishing security situation, and of the difficulties in engaging girls who had already married and dropped out of school:

> Yes, mothers' forum does provide as much support as is possible to send girls to school and address their problems. The only task they cannot do is financial support. Or those families who marry their daughters, they cannot stop them. Because families don't allow their brides to attend school.

Caregiver, Hirat
Endline Evaluation Report, BRAC (Afghanistan)

Adult literacy classes are helping to improve caregivers' perceptions of the value of girls' education. In AKF (Afghanistan) project areas, literacy classes were only implemented from year four (2016), reflecting a change from activities implemented at midline. Although community members have only been enrolled for a short period of time, there is encouraging qualitative evidence that these are effective in improving their perceptions of the value of education:

> *I have participated in the workshop that was about writing and reading. That workshop was very useful for me, in which I learned writing and reading. Unfortunately now it is difficult for me to become literate, but I try my best for my children to become literate so that they become educated.*

Caregiver, Baharak District, CBE community

There is encouraging evidence that literacy classes can be effective in galvanising caregivers to support their daughters' education (see also Marphatia and Moussie, 2013[84]).

CARE (Somalia), which provided evening literacy courses to mothers, found that the classes were having a positive effect on the perceptions of girls themselves, with qualitative evidence suggesting that they now felt motivated to go to school after seeing their mothers participate in the classes. The report suggested that improvements in caregiver literacy were helping with girls' literacy, with girls claiming that mothers were now able to assist them in completing their homework. However, primary data from the girls themselves was not presented in their report, nor did the report present any quantitative data linking the effectiveness of this intervention to improvements in attendance and literacy outcomes.

Community-level barriers

Projects mobilised community support through various means, including through direct interventions — such as community workshops and Community Education Committees — as well as indirect interventions, including involving community members in the management of school and programme activities.

BRAC (Afghanistan) supported the implementation of community workshops at government schools (hub schools), which aimed to raise awareness about how to reduce barriers to girls' education. Evidence presented in their endline report suggests that community attitudes are steadily improving, with 72 per cent of CBGS girls' female carers indicating that community members have become more encouraging of girls to succeed in school in the last 3 years, as compared to 47 per cent at midline.

Qualitative data from BRAC (Afghanistan) and ACTED (Afghanistan) found that **community workshops were pivotal in dispelling myths related to girls' education and educating the community in the benefits of education.** Qualitative reports from ACTED found that their continued presence in the area was working to improve ACTED's reputation among community members, and was helping to dispel myths related to the purpose of the project and improve their support towards girls' education.

At endline there is also **emerging evidence that community members are playing an active role in mobilising support for girls' education.** In AKF (Afghanistan), communities were mobilised to provide in-kind contributions for girls' education as part of their sustainability plan. The project found that the total value of in-kind contribution increased since midline and vastly exceeded their overall endline target, with qualitative findings pointing to widespread support among community members for girls' education. Qualitative findings from ACTED's (Afghanistan) report also found that regular discussions with community leaders helped convince parents of the benefits of girls' education, enlisting their support to send their girl to ACTED courses.

---

[84] Marphatia A, Moussie R (2013) A question of gender justice: exploring the linkages between women's unpaid care work, education and gender equality. International Journal of Educational Development 33(6): 585-594.

In a similar vein, CfBT (Kenya) found that community members and *Community Health Workers* were playing an active role in visiting parents and speaking to them about the importance of education, with some links to improved enrolment[85]:

> *We also take it upon ourselves to visit the household and talk to the parents asking them why their daughter is not going to school. We then advise the parent to speak to their child on the importance of education. If we notice no change we then inform the village head who also visits the household together with a teacher from the school for appropriate action.*
>
> Community member, County 66
> Endline evaluation report – CfBT (Kenya)

In RI (Somalia), the percentage of caregivers and girls who have been exposed to project awareness-raising activities and who report that their perceptions of the importance of girls' education have changed increased from midline to endline across all provinces served by RI (Somalia) (Somaliland: 80% / 96%; Puntland: 73 % / 81%; and Benadir & Galmudug: 82% / 95%). Qualitative data found evidence of communities and mothers taking responsibility for challenging wider beliefs about girls' education within their communities, in addition to reports of Child Education Committee (CEC) members and teachers taking it upon themselves to find a solution to support out-of-school girls' return to school.

There is some evidence that radio messaging on girls' education can be effective in improving enrolment and attendance, although the observed effect can be dampened by spillover effects. In RI (Somalia), 90 per cent of caregivers at endline exposed to radio messaging indicated that their perception of the importance of girls' education had changed, with positive correlations found between girls' enrolment and households with caregivers who had seen or heard an awareness raising message in the past year (see also BBC Media Action's (2017)[86] analysis of the effects of the 'Our School' programme in South Sudan). However, while more caregivers were able to access and listen to the project radio broadcast in urban areas, such as Benadir and Galmudug (where 81% of caregivers heard a radio message on girls' education), access to radios deteriorated in rural areas, such Puntland (50%) and Somaliland (16%).

Similar issues emerged in ACTED and AKF (Afghanistan). In AKF, although there was anecdotal evidence of parents who had decided to enrol their daughters in school after listening to the radio shows, there was also evidence that community members could not always recall the name of the radio shows they had heard. Concerns regarding spill-over effects were also noted by Plan (Sierra Leone), although the focus of their radio intervention was largely to provide children with alternative education and to disseminate sexual and reproductive health rights information, rather than focus on messaging around girls' education.

Girls' aspirations

Most SCW projects address girls' aspirations as barriers to education at endline, which they sought to alleviate through interventions including mentoring, the establishment of girls' clubs and safe spaces, and programmes aimed at improving life-skills and decision-making capacities.

At endline, around half of all projects reported changes in these factors, including in the confidence and esteem of girls (four projects) and the presence of role models (two projects). One project (Plan (Sierra Leone)) reported an increase in the prevalence of early marriage, although it largely attributed this to the outbreak of Ebola at the time, something that has been found in other studies (UNFPA, 2015[87]).

Across the window, there is some evidence that channels used to support girls with their learning directly may be linked to improved attendance and learning outcomes (four projects).

---

[85] According to their endline evaluation report, 13,225 out-of-school girls were identified through community mapping and 9,426 were supported to enrol/re-enter school, exceeding their output-level targets in both instances (8000 and 6000 respectively). In addition, the project exceeded its output-level target for the number of households with marginalised girls visited by Community Health Workers (15,757 against an endline target of 12,000) and broadly met their target for the number of those households accessing health and social support services that facilitate their attendance (4,826 against a target of 5,000).

[86] BBC Media Action (2017). How is radio helping to improve girls' education in South Sudan? Available at: http://downloads.bbc.co.uk/mediaaction/pdf/research-summaries/girls-education-south-sudan.pdf. Accessed 19th September 2017.

[87] UNFPA (2015) 'Rapid Assessment of Pregnant Adolescent Girls in Sierra Leone,' Freetown: UNFPA

> **Key findings**
>
> *Motivation and aspirations*
>
> - Mentoring and girls' clubs are effective vehicles in improving girls' self-confidence at school and their aspirations.
>
> *Extra-curricular activities*
>
> - Extra-curricular activities that directly support girls' learning (e.g. help with reading or homework) are helping to increase motivation, with some evidence that these channels may be linked to improved attendance and learning outcomes.
>
> - Unless support is directly related to learning, it is unclear whether addressing girls' low aspirations and lack of self-confidence is linked to improved learning.
>
> *Early marriage and pregnancy*
>
> - Early marriage and pregnancies remain a key barrier to education for girls after puberty.

Lack of self-confidence

Projects sought to address girls' aspiration and decision-making in multiple ways, although they were not always explicit about how these interventions were intended to improve educational outcomes. BRAC (Afghanistan), for example, devised a mentoring programme to provide a safe space for girls and increase their self-esteem. Similarly, WV (Zimbabwe) established *Power Within (PW)* clubs with the aim of providing an enabling environment and fostering leadership skills and knowledge of girls' rights. However, across the window there was limited evidence of a clear theory of change linking improved empowerment to educational outcomes.

At endline, we find that girls' levels of self-confidence are improving as a result of mentoring, girls' clubs and life skills interventions. Camfed (Zimbabwe and Tanzania), for example, delivered the *My Better World (MBW)* programme, which comprised of books as part of an 18-month long programme aimed at improving life skills and decision-making. Students' participation in this programme was significantly associated with increases in students' enjoyment of school and confidence in their academic progress over time. Evidence from BRAC (Afghanistan), RI (Somalia) and WV (Zimbabwe) also found improvements in girls' self-confidence, motivation and/or empowerment.

Girls' clubs are improving girls' aspirations and their sense of school belonging. Girls' clubs delivered by STC (Mozambique) were significantly linked to *school belonging*[88] with members exhibiting higher levels of *school belonging* (4.10) than non-members (3.92). Camfed (Zimbabwe and Tanzania) also found that girls were significantly more likely to say that the *My Better World* sessions had made them feel more integrated in society and more confident about the future than boys. RI (Somalia) and WV (Zimbabwe) both presented qualitative evidence linking girls' clubs to improvements in their aspirations and in the relevance of their education in later life. **None of these projects presented conclusive evidence linking these particular effects to improved educational outcomes.** This finding is reflected in the academic literature where there is considerable evidence of a range beneficial outcomes from girls' clubs such as increased empowerment, strengthened livelihoods, etc., (Marcus and Brodbeck, 2015[89]; Unterhalter et al, 2014[90]), but little evidence of improved learning outcomes, at least on account of the intervention of girls' clubs alone (Buchmann et al, 2017[91]).

Girls' clubs and mentors also appear effective in empowering girls to report safety issues. WUSC's (Kenya) qualitative evidence found that girls' clubs empowered girls to speak up against inappropriate sexual contact:

---

[88] This is an index measured by the project

[89] Marcus and Brodbeck. (2015). Girls' clubs and empowerment programmes. ODI Research and Practice note, ODI, UK.

[90] Unterhalter E, North A, Arnot M, Lloyd C, Moletsane L, Murphy-Graham E, Parkes J, Saito M (2014) Interventions to enhance girls' education and gender equality. Education Rigorous Literature Review. Department for International Development

[91] Nina Buchmann, Erica Field, Rachel Glennerster, Shahana Nazneen, Svetlana Pimkina, Iman Sen. (2017). Power vs Money: Alternative Approaches to Reducing Child Marriage in Bangladesh, a Randomized Control Trial. JPAL working paper

> *I am in the Sexual Gender Based Violence club where we talk about healthy relationships between us and the teachers and also our parents and ourselves as pupils. … This club has helped end romantic relationships between teachers and girls because girls are now empowered and report those involved. We also report those boys and girls who spend too much time together engaging in inappropriate touch.*

Girl student, Turkana County
Endline Evaluation Report, WUSC (Kenya)

Similarly, WV (Zimbabwe) cited evidence from mothers who felt that the clubs had given the girls more confidence in discussing their sexual well-being and improving communication with them as a result. RI (Somalia) found that relationships between Community Education Committees (CECs) and girls' clubs strengthened during the course of the intervention, with girls reportedly sharing their needs and concerns openly with female members of the CEC.

There is encouraging evidence that increasing academic confidence and enjoyment in education through extra-curricular activities may be linked to improved attendance outcomes. Two projects, BRAC (Afghanistan) and WV (Zimbabwe), found evidence of positive links to attendance for girls taking part in Girls' Clubs or working with mentors, as discussed below.

In BRAC (Afghanistan), the mentoring scheme helped to offset some of the broader downwards trends in attendance. The proportion of girls with mentors claiming to never miss school fell from 71 per cent at midline to 63 per cent at endline, while for girls without mentors this fell from 91 per cent to 64 per cent at endline. Qualitative evidence indicates that this may be linked to girls' increased academic confidence and motivation as a result of interacting with their mentors:

> *I had difficulties in learning mathematics in the past year, I got 16 points in it, but this year I got 26 points. My mentor helps me with learning mathematics. She has a positive effect on us…I had problems in learning mathematics and I even didn't want to come to school, but now that I have learned I come every day to school and I am not absent much and I am so happy with my mentor.*

Girl student, Kabul
Endline Evaluation Report, BRAC (Afghanistan)

In WV (Zimbabwe), girls who joined the *Power Within* club were significantly more likely to be enrolled and have higher attendance at endline compared to girls who did not. In addition, girls who *attended* a school with a PW club were found to be significantly more likely to have higher EGRA and EGMA scores at midline, with girls who were actually *part* of the PW club more likely to score higher in these at both midline and endline than girls who did not join a PW club.

In qualitative findings, girls also spoke of using books from the *Happy Readers* programme during the clubs, with *Power Within* participants found to have improved their reading fluency 18 per cent above their control peers. Anecdotal evidence in their report attributed these to improvements in self-confidence in academic ability and to the increased exposure to books via the *Happy Readers* programme.

There is some evidence to suggest therefore that **supporting girls directly with their learning through extra-curricular and non-formal education can lead to improved learning outcomes.** Qualitative evidence from projects' endline reports (STC (Mozambique), BRAC (Afghanistan), and WV (Zimbabwe)) attribute these improvements in girls' learning to increased self-confidence and motivation. Some projects also found evidence of a positive link between these interventions and attendance outcomes (WV and BRAC).

In BRAC (Afghanistan), EGRA and EGMA scores for girls with mentors increased from midline to endline, with significant effects over and above the scores for girls without mentors[92]. Qualitative evidence from their endline report found that the 'main' role of mentors was in coaching and supporting girls' learning, including assisting girls with their homework and helping them study for their forthcoming exams. Narratives from mentees suggested that

---

[92] For example, girls with mentors in Grade 6 improved their EGRA scores from 65.33 at midline, to 107.31 at endline, while girls without mentors had more modest improvements from 88.85 at midline to 101.77 at endline. A similar pattern emerges in terms of improvements in EGMA scores and for girls in other classes.

such direct help with their learning had increased their motivation in school, since it improved confidence in their academic ability.

Evidence from STC's (Mozambique) project similarly found that improvements in self-efficacy led to increased motivation and better learning outcomes. Due to missing evidence, it is unclear how the Girls' Clubs increased motivation. However, from the theory of change, these clubs were intended as a vehicle through which girls are assigned peer educators who act as their mentors in academic and pastoral issues. Quantitative evidence from their endline report found that girls' club membership affects both literacy and numeracy at endline with significant effects.

Box 12: Academic confidence – what has changed overall?

Findings presented below are based on primary data gathered by the EM at endline. Treatment averages are shown on top and control averages are shown on the bottom. Statistically significant changes are indicated in the text below.
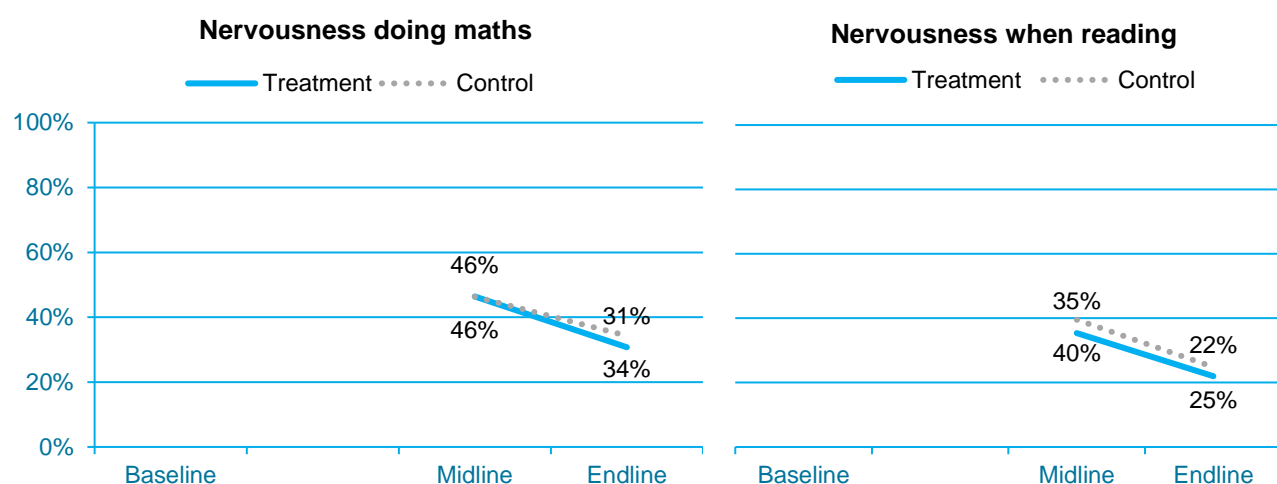
Household Survey (Evaluation Manager) | "I get nervous when I have to do maths in front of others". Proportion of caregivers answering "yes". Average across all SCW EM project areas, excluding Camfed (Zimbabwe and Tanzania).

Household Survey (Evaluation Manager) | "I get nervous when I have to read in front of others". Proportion of caregivers answering "yes". Average across all SCW EM project areas, excluding Camfed (Zimbabwe and Tanzania).

Data gathered by the EM confirm projects' reported findings at endline, which suggest that girls' confidence in maths and reading is improving at the window level. **Interventions seem to be especially effective in improving girls' academic confidence in maths**, with the share of households in treatment areas reporting their girls as feeling nervous doing maths in front of others falling from 46 per cent at midline to 31 per cent at endline, as compared with a decrease in control areas from 46 per cent to 34 per cent respectively. This effect is statistically significant at the 10 per cent level for girls in treatment areas.

Early marriage and pregnancy

None of the projects presented conclusive evidence at endline about improvements in early marriage and pregnancy, due in part to the difficulty of generating accurate data (as early marriage is often illegal). According to the 'Economic impacts of child marriage' study (ICRW/World Bank, 2017[93]) the global costs of early marriage are likely be in trillions of dollars, in part due to their effect on education.

Projects have had limited reported success in reducing the effect of early marriage and pregnancies, with no conclusive effects on learning outcomes. BRAC (Afghanistan), for example, presented qualitative evidence suggesting that interventions such as Mothers' Clubs and the involvement of *shuras* have not been particularly effective, since early marriage was often perceived as outside of their direct control. STC (Ethiopia), who worked with stakeholders to promote Family Laws against early marriage and FGM, found from qualitative evidence that

---

[93] ICRW/World Bank (2017). Economic Impacts of Child Marriage: Global Synthesis Report.

girls who rejected *absuma*[94] marriage in favour of attending school sometimes faced negative consequences, including being required to pay compensation. WUSC (Kenya) also found that perceptions of girls who complete primary education as less "marriageable" still largely prevailed at endline, as they are perceived to be older and more independent.

Early marriages and pregnancies remain among the biggest barriers to girls' retention in school. Findings from the Girls' Education Challenge Thematic Research Report (2016: 24, 34) point to the prevalence of early marriage among girls as a household livelihood strategy, with parents either 'marrying girls off' or girls choosing to marry early to escape the poverty of their parents' household[95]. At endline, both STC (Mozambique) and CARE (Somalia) found that teachers most frequently cited marriage as the main barrier to girls' education.

Camfed (Zimbabwe and Tanzania) sought to address the prevalence of early marriage and pregnancy through a holistic approach, including the provision of in-kind support (through the *Step Up Fund*) as well as wrap around support from local authorities and parent support groups and a life skills programme (*My Better World*) intended to improve girls' aspirations for the future. Even with SUF support, however, survey results at endline found that marriage and pregnancy were among the most common reasons for drop-out among marginalised girls receiving Camfed support, with 65 per cent of drop-outs attributed to marriage in Zimbabwe, and 73 per cent of drop-outs attributed to pregnancy in Tanzania.

Although Plan (Sierra Leone) aimed to reduce the effect of early marriage and pregnancy through radio broadcasts, evidence presented in their endline report suggests that only 35 per cent of girls in the sample were exposed to these. The problem of reaching women and girls via radio broadcasts has been observed in other studies (e.g. Doherty and O'Connor, 2017[96]) as even when the household has a radio, male household members may take it to the field with them. In addition to this, contextual factors such as the Ebola epidemic were attributed to a spike in early marriage and pregnancy, with qualitative findings suggesting that acute levels of poverty brought on by the epidemic created incentives for girls to marry early rather than attend school.

Early marriage was a central component of Child Hope's (Ethiopia) intervention, which developed a series of demand and supply-side interventions to improve the school environment and increase the premium attached to girls' education. Given the holistic nature of the project in addressing girls prone to early marriage, limited data was available from their report in terms of attributing changes in girls' ability to remain and perform well in school to particular activities or interventions.

Violence

Most projects identified violence and safety issues as barriers to girls' education, such as sexual harassment and sexual abuse (both on the way to school but also from adults in charge) and the use of physical punishment and discipline by teachers in class. Some projects sought to address these barriers as a core part of their overall design, through building classrooms closer to school or opening girls-only schools, while others carried out more discrete interventions, including implementing Child Protection Policies, strengthening referral and complaints mechanisms, and training teachers on child protection.

Evidence about projects' effects on violence and safety was largely inconclusive, with only half of all projects that sought to address these barriers (six out of 12 projects) presenting any data on this barrier at endline. Two projects reported improvements in the incidence of in-school violence, with the same number reporting a worsening of fears of violence. Two projects observed changes in the use of corporal punishment at endline, although there was conflicting evidence on the direction of this change.

---

[94] In the Absuma culture of Afar, marriage is a social contract between two exogamous tribes who are forced to exchange women between themselves (definition taken from the STC (Ethiopia) endline evaluation report).
[95] Evaluation Manager Girls' Education Challenge Fund (2016). Narrow windows and Revolving Doors. GEC Thematic Research Report, March.
[96] Trish Doherty and Karen O'Connor (2017). Improving maternal and child health through media in South Sudan: Final Evaluation. BBC Media Action, BBC, UK.

> **Key findings**
>
> - Multi-dimensional interventions are improving the response to in-school violence, although these have had limited success in allaying girls' fears of violence at school.
>
> - While teachers are becoming increasingly aware of what constitutes abuse or violence, there are prevailing barriers and issues around reporting violence, particularly among men.
>
> - Although there is some evidence that teachers are making increasing use of non-violent disciplinary measures, evidence is mixed in relation to the use of corporal punishment.
>
> - Evidence is inconclusive on the effect of violence interventions on attendance and learning outcomes.

In-school violence & fears of violence

There is some evidence that multi-dimensional interventions are improving the response to in-school violence. Both Camfed (Zimbabwe and Tanzania) and WV (Zimbabwe) identified sexual abuse and sexual violence as barriers to girls' education, particularly within schools. Qualitative research from Camfed and World Vision found that stakeholders believed the interventions had strengthened the whole child protection process:

> *There is a case of a child who was being abused by her uncle at home. The uncle wanted to sexually abuse the girl. This came to light when we got an anonymous tip off from the suggestion box. Our Child Protection Committee engaged the family of the child and warned the perpetrator who took heed of the warning…There is one particular case of a child who was in grade seven who was given money by some boys so that they could get sexual favours. Fortunately, this came to the attention of some girls who are members of Power Within who notified some elders. The case was dealt with amicably.*
> Caregiver, Lupane
> Endline Evaluation Report, WV (Zimbabwe)

The example above highlights how the multiple interventions (suggestion boxes, Child Protection Committees and Power Within clubs) were working together effectively to strengthen the reporting process and resolve the instance of child abuse.

In their endline evaluation report, Camfed (Zimbabwe and Tanzania) found that significantly more intervention schools had responded to abuse cases in the past 12 months than in comparison schools. Nevertheless, Head Teachers in intervention schools were significantly less likely than their counterparts in comparison schools to believe that incidents would be reported. Further evidence from their endline evaluation report suggests this may be because teachers are becoming more aware of what constitutes abuse or violence, rather than due to a lack of confidence in the reporting system.

Two projects found evidence of issues around reporting violence, particularly among men, with changes in norms likely to be a slow process. In Camfed's project areas, there were no significant differences in reported awareness of Child Protection Policies between baseline and endline, with school visits finding that the prevalence of violence was greater than was being reported due to an 'acceptance of abuse as a longstanding practice, especially against girls (Endline Evaluation Report, Camfed (Zimbabwe and Tanzania): 101)[97]. The endline report found that there were particular concerns regarding underreporting of abuse perpetrated by males, which is echoed by evidence reported by World Vision regarding the reluctance of men to report abuse:

> *A Binga Male Champion spoke about the challenges of trying to reduce or eliminate gender-based violence in his community when he said, "most men are reluctant to report cases of sexual abuse particularly when this is happening within the family. Women are forthcoming in terms of reporting cases of abuse because they confide in their colleagues who then report to the responsible authorities… I think this is because of lack of awareness among men and the fact that most of them are the perpetrators."*
>
> Endline Evaluation Report – WV (Zimbabwe)

---

[97] CIDT (2017) Endline Qualitative Study of the Camfed Girls' Education Challenge Step Change Window (GEC SCW) funded programmes in Tanzania and Zimbabwe, pp.67 in Annex 11 cited in Endline Evaluation Report – Camfed (Zimbabwe and Tanzania), p101.

This highlights that interventions such as suggestion boxes were not necessarily viewed by men to be relevant to them, and that such interventions are likely to require a long time before changing norms around gender-based violence.

Efforts to improve child protection policies and practices have had limited success in allaying girls' fears and concerns about violence. As part of its theory of change, CARE (Somalia) aimed to deliver child-friendly training to teachers in an effort to encourage more regular attendance. Nevertheless, 19 per cent of girls at endline reported feeling afraid at school, as compared to 11 per cent at midline; a finding it attributes to external factors. However, data on this is inconclusive as it is not possible to ascertain the cause of girls' fears.

In Camfed (Zimbabwe and Tanzania), despite endline research finding that the majority of schools had child protection policies, perceptions of safety among girls have broadly remained the same since midline, with the percentage of students who believe that most or all incidents of child abuse are reported falling short of their endline targets in both Tanzania and Zimbabwe. Qualitative research highlighted lingering concerns among girls concerning teacher bullying, as well as peer harassment and bullying from peers, both within and on the way to school. **This suggests that child protection policies and associated training addressing violence may not be effective (at least in the short-term) in alleviating fears of violence, since these are often too indirect to be perceived by the girls themselves and would not address violence occurring on the way to school.**

BRAC (Afghanistan) provided student organisers (*khalas*) to accompany government school girls to and from school. In their endline evaluation report, however, BRAC found that only two girls out of 20 indicated that this was their method of travel into school. Yet, insecurity appears to be a growing problem, with the proportion of families stopping girls from going to school due to insecurity rising from 52 per cent at midline to 85 per cent at endline. This could be because girls and households are not aware of *khalas'* role, or that they are not perceived to be effective in keeping girls safe from external security factors. Indeed, qualitative evidence found that girls perceived *khalas'* main role to be in delivering stipend payments, with some girls reporting that they tend to be accompanied to school by other family members or other girls from the same school.

Child Hope (Ethiopia) provided letter link boxes in schools to enable children (and in particular girls) to report issues and concerns anonymously, with qualitative evidence suggesting that they were also being used as a forum to air their views and communicate with peers on school and personal issues. This could explain the association found at endline between schools with *Letter Links* and self-esteem scores, as the evidence suggests that girls felt empowered to raise their voice, particularly when they felt their views were being listened to.

Looking more closely at their endline report, we find, however, that the incidence of violence has not changed since midline for either treatment or control groups, suggesting that while providing a channel for girls to directly report incidents may be effective in improving girls' sense of safety, this may not necessarily be effective in reducing the incidence of physical violence.

Corporal punishment

Another form of violence that affects girls' ability to enrol, attend and learn in school is the use of corporal punishment, which could affect motivation to attend school and concentrate in class. This finding is supported by evidence from other studies such as Young Lives (Pells and Ogando, 2015[98]) and reviews such as Gershoff (2017)[99].

At endline, evidence is inconclusive in relation to the use of corporal punishment. AKF (Afghanistan) provided child protection training to teachers as part of a wider series of interventions addressing the safety of the learning environment. At endline, the proportion of government school classes observed to be free of violence increasing from 80% at midline to 91.5% at endline[100]. Moreover, a significant positive correlation was found at endline between the absence of corporal punishment administered by teachers, and girls' literacy and numeracy

---

[98] Kirrily Pells, María José Ogando. (2015). Corporal punishment in schools. UNICEF Office of Research – Innocenti Discussion Paper 2015-02.

[99] Gershoff, E.T. (2017). School corporal punishment in global perspective: prevalence, outcomes, and efforts at intervention. Psychology, Health & Medicine Vol. 22 , Iss. Sup 1.

[100] While the proportion of CBE school classes observed to be free of violence decreased slightly from 100% to 96.9%, the project attributes this to fewer cases of missing data at endline rather than an increase in the use of violence.

scores, as well as with girls' attendance reported by their carers. This could be linked to improvements in confidence and motivation, with qualitative data from community members suggesting that the use of corporal punishment tended to negatively impact students' academic confidence and motivation to study.

CARE (Somalia), which also delivered child protection and child-friendly training, found in classroom observations however that 70 per cent of teachers were using physical punishment in the class, up from 16 per cent (although this diminished overall from 96 per cent of teachers at baseline). However, there are serious limitations with the reliability of the data, with the project noting that that these figures may be inflated due to sampling biases.

What other non-GEC activities and external events have happened, and with what effect on barriers and intervention effectiveness?

At endline, projects reported on a number of external events and factors that occurred during the course of project implementation, which may have directly or indirectly affected GEC project delivery, such as political unrest, policy changes, and spillover effects from other non-GEC interventions.

Violence and political unrest threatened the delivery of some GEC projects

Around half of projects (six projects) cited **violence and political unrest as barriers to project delivery**, leading to some GEC activities being temporarily halted during the course of the intervention or altogether stopped. According to a UNESCO report, conflict and unrest tends to reinforce gender inequality, since displacement often results in families resorting to 'coping mechanisms that disadvantage girls, including child domestic work and child marriage' (UNESCO 2016: 22)[101].

Both BRAC (Afghanistan) and ACTED (Afghanistan) reported increased threats of violence posed by conflict, leading to the temporary closure of CBGS classes in BRAC areas in Kunduz. Although the schools are said to have reopened once security had been restored, it is not known how long the classes were closed for.

In STC (Mozambique), military and political tension in Tete meant that certain schools and communities in Guro and Barue were inaccessible for parts of 2015 and 2016, while in both STC (Mozambique) and CfBT (Kenya) project areas, concerns for safety interrupted activities such as community conversations and literacy promotion. In Child Hope (Ethiopia), there were reports of increased political insecurity and community violence, although the effects of this on the intervention are not clear.

WUSC (Kenya) reported attacks by Al-Shabab that were said to have led to important policy developments during the project's implementation, including the repatriation of Somali refugees. The political uncertainty for refugees was found to have negatively impacted student attendance in Dadaab's schools in 2016, while security concerns were cited as a key reason for teacher absenteeism at both baseline and midline.

Spillover effects from non-GEC interventions may have impacted the effectiveness of some GEC projects

Six projects reported other non-GEC interventions active in their treatment or control areas, with mixed effects on GEC outcomes.

In their endline report, CfBT (Kenya) reported that the TUSOME national programme jointly funded by DfID and USAID (which focused on early grades) had raised questions of waste of resources, prompting the project to focus on upper grades instead. Nevertheless, CfBT notes that they had already procured learning materials for some early grades, leading to duplication of grade two literacy materials in project schools.

WV (Zimbabwe) noted challenges enlisting support from the community as they were not offering immediate financial incentives unlike other NGOs operating within their areas. WV also noted that negative experiences of community members with other projects affected their willingness to engage with WV, although it is not clear what the impact of this was on the project's ability to support girls' attendance and learning.

BRAC (Afghanistan), Plan (Sierra Leone) and Child Hope (Ethiopia) all found evidence of **similar interventions working in their areas that may have contributed to improvements in attendance and learning outcomes.** STC (Ethiopia) reported that the World Food Programme School Feeding intervention had been active since 2015, with a component providing water and school materials in some areas. Although they claim that this led to

---

[101] UNESCO (2016). *Gender Review: Creating Sustainable Futures for All*. Global Education Monitoring Report. Paris, France: UNESCO.

improvements in enrolment and attendance outcomes, they also suggest that none of the attendance data was based on schools that were targeted by the programme. In Plan (Sierra Leone) project areas, other organisations such as the Global Partnership for Education and UNICEF were said to be supporting radio-learning programmes in the wake of the Ebola outbreak, while in Child Hope (Ethiopia), NGOs such as *Good Neighbours from Korea* were reported to be providing education materials and school uniforms to children in the area.

Natural disasters, health epidemics and other economic pressures all affected families' ability to afford the cost of schooling

Drought negatively affected the implementation of five projects (STC (Ethiopia), Camfed (Zimbabwe), WV (Zimbabwe), Child Hope (Ethiopia) and RI (Somalia)). STC (Ethiopia) reported that the 2015 drought led to delays in school infrastructure improvements, with certain activities at the community-level (e.g. VSLAs) also difficult to support. Child Hope (Ethiopia) reported that drought conditions led to a lack of access to water at school, with teachers reportedly unwilling to work at school, and girls interrupting school to travel home to drink water.

RI and Camfed both reported reduced levels of household income and food security among their target communities, making it more difficult for households to prioritise spending on girls' education. This in turn affects girls' ability to attend school as they face increased pressure to marry early or support their families with income generation.

As mentioned elsewhere in the report, **the outbreak of Ebola** had significant implications for the delivery of Plan's (Sierra Leone) project, with schools closed for about nine months.

A favourable policy environment enabled project delivery for some projects, although certain policies did not always have the intended effects

A handful of projects cited **policy developments and positive relations with the government as factors facilitating the implementation of project activities.** WV (Zimbabwe), CfBT (Kenya) and Camfed (Zimbabwe and Tanzania) indicated strong institutional commitments and partnerships with government stakeholders, which enabled the projects to progress with their activities. WV (Zimbabwe) also welcomed the government's School Feeding Program initiative in 2016 targeting lower grades. This is likely to have had a significant effect on attendance and learning outcomes, although detail on the number of schools affected and the corresponding breakdown by treatment and control groups was not available in their report.

Nevertheless, certain government policies did not always have the effects anticipated by projects. In Tanzania, school fees were abolished in 2015, which led to a reported spike in enrolment among schools in Camfed's (Tanzania and Zimbabwe) project areas. In the short-term, Camfed found that this negatively impacted the quality of education being delivered, but also led to a reduced willingness of parents to pay for school expenses, as they perceived the government to now be in a position to extend their financial support to meet other school costs.

RI (Somalia) pointed to the **low capacity of the state to support or fund public services** as an obstacle to ensuring the sustainability of project outcomes. In particular, there were reports of teachers not being paid on time and a lack of financial and material support from ministries to schools, which resulted in slow progress towards ensuring the **sustainability** of interventions.

Teacher strikes and shortages impacted the delivery of some projects

Both CfBT (Kenya) and WUSC (Kenya) reported teacher strikes during project implementation lasting between one and several months. In WUSC's (Kenya) project areas at least, this was reported to have negatively impacted attendance rates.

According to WV's (Zimbabwe) endline report, changes in education policy in Zimbabwe meant that non-qualified teachers were no longer allowed to teach, the effect being that they were removed from target schools, resulting in increased class sizes and reduced quality of instruction. Certain interventions such as *Power Within* clubs and the *Happy Readers* programme were said to be particularly affected due to a lack of teachers to implement the activities.

What were the unintended consequences of SCW interventions?

In addition to reporting on external events and factors that occurred during the course of project implementation, at endline projects also reported on a range of unintended effects (both positive and negative) that occurred as a result of SCW interventions.

There was evidence of resentment and jealousy among excluded boys, although projects are beginning to address this in their interventions

A number of projects (five projects) reported instances of **jealousy and resentment** among boys. IRC (DRC), Plan (Sierra Leone), WUSC (Kenya), Child Hope (Ethiopia), STC (Ethiopia) and CARE (Somalia) all presented examples of boys complaining about feeling excluded or jealous for not being targeted by the project, although some projects took measures to mitigate or address these.

IRC (DRC) highlighted instances of boys teasing girls for being project beneficiaries (especially those receiving bursaries), with reports of girls being beaten or targeted for sexual harassment in the most severe cases. Similarly, WUSC (Kenya) spoke of boys disrupting remedial classes as a result of feeling excluded from the programme. Other examples emerged in Child Hope's (Ethiopia) and Plan's (Sierra Leone) projects, with one boy in Kono stating: "*There is no BEC, only GEC […] we would like to be receiving bursaries too*" (Endline Evaluation Report, Plan (Sierra Leone)).

Following complaints received at midline, STC (Ethiopia) provided materials for both boys and girls, which helped mitigate further concerns and feelings of exclusion among boys. In Camfed (Zimbabwe and Tanzania), the project credited its ability to mitigate risks of backlash from men and boys to involving stakeholders regularly in the design and progression of the intervention. Although Child Hope's (Ethiopia) report described instances of boys teasing girls who benefited from the project, the report acknowledged that this had reduced since midline due to wider efforts undertaken by the project to engage with the needs of boys, including through the facilitation of *Good Brothers'* Clubs, as well as the provision of *Letter Links Boxes* to accommodate their concerns.

Some interventions improved relationships between boys and girls

Nevertheless, two projects found examples of **increased support among boys for girls' education**. Plan (Sierra Leone), reported that boys were feeling pleased that girls were being targeted and empowered through the intervention to remain in school.

At endline, AKF (Afghanistan) conducted FGDs with boys to ensure their voices were captured during the intervention. Boys in the groups were reported to express support for their sisters to go to school, including by taking on additional household labour to enable girls to attend school. Less positive examples, however, emerged of boys remaining out of school to do household labour, which the report acknowledged as among the negative consequences of the intervention.

Jealousy and backlash reported from control schools and communities not targeted by the intervention

Child Hope (Ethiopia) and RI (Somalia) both found instances of hostility from communities not receiving project support, with RI indicating conflicts escalating between communities over access to resources.

At the household-level, there were **mixed effects on relations between household members**. WV (Zimbabwe) reported that relations between men and women had improved due to increased household income attributed to the VSLA scheme. RI (Somalia), however, reported tensions within the household due to the increased burden on families to provide school-related materials for their enrolled children. RI and CARE (Somalia) both reported a sense of perceived injustice within families since not all siblings were able to benefit from the projects' activities.

There were some reports of disappointment and confusion regarding the beneficiary selection process

Another unexpected consequence to emerge was the **raising of expectations and disappointment among some project beneficiaries.** Some community members in AKF's (Afghanistan) project areas reported feeling disheartened by the one-time enrolment practice, with added disappointment that boys were not benefiting as much as girls in CBE communities. The project reported that this resulted in poorer perceptions of education and of the value of girls' education, particularly among men. Similarly, WUSC (Kenya) described how a shortage of space

in secondary schools led to negative attitudes from community members in terms of the perceived value of girls' education.

Reports from BRAC (Afghanistan) and Plan (Sierra Leone) identified concerns regarding the selection of beneficiaries. Qualitative research with Plan's stakeholders reported that if a girl was absent on the day that the implementing partners chose to visit the school, they would not be selected to benefit from the programme. Girls not chosen by the intervention were said to feel **jealous and resentful towards beneficiaries,** as well as confused by the selection process.

Isolated reports of resources being misused in certain project areas

Two projects presented evidence of resources being misused, although it is important to note that these emerged as isolated examples within the project reports and only one of them relates to misappropriation. CfBT (Kenya) for example reported instances of girls donating their solar lamps to neighbouring hospitals to provide lighting for mothers going into labour. CARE (Somalia) reported an instance of a teacher withholding school resources provided by the project (such as textbooks) and selling them for a profit as part of his own business. It should be noted that the FM's whistleblowing process and in-country project monitoring would provide a more systematic and comprehensive source of information about the misuse of resources across the GEC.

Unintended positive effects on girls' aspirations around early marriage and pregnancy

Three projects reported unintended effects on girls' aspirations around early marriage and pregnancy. Evidence presented by WUSC (Kenya), RI (Somalia) and CARE (Somalia) found that girls' aspirations are improving, with girls speaking of their rights in relation to choosing their own spouses and in delaying marriage to pursue their own educational and career aspirations.

WUSC (Kenya) followed up on girls who dropped out of school due to pregnancy and found that women in Kakuma camps who had given birth were reported to have returned to school to complete their education once suitable care for their children had been identified. In their endline report, CARE (Somalia) suggested that as girls were progressively exposed to different role models during the course of the intervention, they were happier to delay marriage in order to pursue their own career plans.

At endline, STC (Ethiopia) reported stories of girls facing complaints from wider community members for rejecting *absuma* marriage in favour of attending school, although the report concludes that ultimately these led to gradual improvements in community attitudes.

Some evidence of increased motivation to attend and perform well in school among scholarship and bursary girls

Plan (Sierra Leone) and WUSC (Kenya) also found evidence of increased motivation among girls receiving scholarships and bursaries in their project areas. According to their endline report, boys in Plan's project areas found that bursaries were motivating girls and encouraging them to perform their best in school.

WUSC provided scholarships to girls, and stakeholders reported that these had a positive effect on their motivation to sit and perform well in their school leaving exams (Kenya Certificate of Primary Education - KCPE). The report found that the number of girls sitting the KCPE and scoring well on the exam had risen among intervention areas since the project started in 2013. However, data on KCPE scores in control areas were not presented in their report.

There is evidence of interventions overlapping and unexpectedly working well with one another in some projects

Both BRAC (Afghanistan) and Camfed (Zimbabwe and Tanzania) found evidence of spillover effects within their interventions, with unexpected benefits from project activities to wider groups of beneficiaries.

BRAC, for example, reported mentors indirectly benefiting from teacher trainings, with mentor girls meeting with teachers to discuss the mentoring programme and benefiting from the knowledge teachers received during the trainings. Camfed reported that study guides were found to have been incorporated into classroom delivery, which had encouraged teachers to make greater use of child-centred teaching pedagogies as girls no longer had to share and copy directly from a single book. In their report, Camfed also noted that Learner Guides had established study groups outside of school hours, which had not been envisaged in the design of the programme. While the project

welcomed the unintended benefit, they cautioned against the risks to girls as they became burdened with school administration requirements.

Interventions are reaching wider groups other than direct beneficiaries, with mixed effects

Three projects presented evidence of interventions reaching girls other than those directly targeted by the projects.

BRAC (Afghanistan) reported stipends benefiting siblings, with beneficiaries using their stipends to cover the costs of stationery and other school supplies of siblings in addition to their own. In their endline reports, BRAC and AKF (Afghanistan) also found evidence of students enrolling in classes they were not eligible for, although it is unclear what the effects of this were.

There were more limited reports of non-beneficiaries 'playing the system' to benefit from GEC interventions. WUSC (Kenya) reported that girls from non-WUSC areas were reported to have moved into project areas in order to access inputs and resources. The project found that increased enrolment was hampering the quality of education due to increased class sizes and limited resources. CfBT (Kenya) found evidence of certain households in Nairobi and Mombasa moving their children out of their existing schools to attend better performing schools, owing to the increased choice afforded by their cash transfer. Following these reports the project amended the conditions associated with the cash transfer to make these semi-conditional on enrolment within the girl's existing school.

Key findings – What has worked, why and with what effects?

At endline, evidence from projects and the effects of particular interventions on girls' attendance and learning was inconclusive in a number of cases (see Table 22 in Annex J). By and large, this was because projects did not present sufficient or reliable evidence linking individual activities with their effects on learning outcomes. Most projects tended to focus evidence in the effectiveness section of their reports at the output or intermediate-outcome level rather than at the outcome level. This was compounded by a lack of clear pathways to change for many projects, which made the question of attribution more difficult. In other cases this was because the type of intervention was broad or overarching, meaning that the effects were diffused and difficult to capture[102]. In other cases still, this was because the quality of evidence presented was poor or missing. On the whole, projects were more successful in identifying the effects of individual interventions on barriers to learning and attendance, rather than the direct effects on attendance and learning outcomes.

Economic interventions and in-kind support helped to reduce the cost of schooling and led to improved attendance among girls. Scholarships, bursaries and stipends all helped to offset the costs of education, and enabled households to overcome difficulties in affording school fees and sending their girls to school, while sanitary wear and menstrual supplies provided by projects have encouraged more regular attendance. Increasing income through loans, savings and cash transfers improved caregivers' willingness towards spending on girls' education, although these did not have any reported effects on girls' learning outcomes.

Infrastructure and facilities improvements have positively impacted on girls' attendance, while learning materials led to improvements in girls' learning. Accelerated learning and alternative schools are helping girls catch up, with positive effects on their learning outcomes, especially in BRAC and AKF (both Afghanistan). Improved teaching methods have helped engage girls in class, although, with the exception of Child Hope (Ethiopia), there is limited evidence that these have improved girls' attendance and learning outcomes. Generally this is due to a reporting deficit; either because projects did not sufficiently measure the effects of these interventions on learning outcomes or because they did not address them in their report.

Community-based activities such as meetings have encouraged caregivers and community members to take an active role in supporting girls' education, leading to improvements in girls' attendance. Radio messaging, community workshops and mothers' groups have all had some success in improving attitudes

---

[102] For example, in the case of building schools nearer girls' homes to improve girls' safety on the way to school and/or to address distance barriers to school. The effects of these are difficult to isolate from those that may occur as a result of other activities organised as part of the intervention, such as improvements to classroom or sanitation infrastructure.

towards girls' education, with encouraging qualitative evidence of community members playing an active role in checking up on girls' absences from school and finding solutions to help them return to school.

Extra-curricular activities have raised girls' self-esteem and confidence in their academic ability, with some improvements in learning outcomes. Mentoring schemes and girls' clubs have been effective in supporting girls with their learning, and have increased their academic confidence, with positive impacts observed on learning outcomes. However, it is unclear whether these improvements in learning are attributable to improved confidence in girls' academic ability, or to the direct support they received with their learning, or to selection effects in relation to which girls participated in the activities.

Interventions addressing violence have not had conclusive effects on girls' attendance or learning outcomes. There is some evidence that the response to in-school violence is improving, but these have not translated into improved confidence that incidents will be reported among teachers, and girls still report being fearful at school.

External factors and unintended consequences

Contextual factors were identified by a number of projects as barriers **inhibiting project delivery.** Political unrest, natural disasters and economic downturns all impacted the implementation of projects, while economic pressure and security threats sometimes forced families to migrate, leading to irregularities in terms of girls' attendance and enrolment.

Nevertheless, several projects cited **favourable policy environments** that facilitated project delivery, including good relationships with government stakeholders which helped project activities progress smoothly.

Since baseline, most projects also identified a number of **unintended effects on non-beneficiaries**, including confusion over the beneficiary selection process and feelings of jealousy and resentment both within families, but also among control schools and neighbouring communities not targeted by the intervention.

Feelings of exclusion and resentment were most pronounced in the case of **boys, with some reports of a negative backlash against beneficiary girls,** including instances of bullying, harassment and verbal abuse reported towards beneficiary girls. Some projects adapted their approach to respond to these concerns, both by including boys' perspectives during their research, but also by implementing activities targeting boys.

Projects also spoke of **unexpected positive effects, including benefits to wider groups** than those directly targeted, and improved relations between men and women. There was also evidence from some projects of beneficiaries indirectly benefiting from overlapping interventions implemented by the project.

Lessons learned

- As at midline, evidence suggests that projects that had **clearly identified the most important barriers in their local contexts were the most effective** in improving girls' learning outcomes– for example, BRAC (Afghanistan), ACTED (Afghanistan) and Camfed (Zimbabwe – Tanzania), who all demonstrated a clear understanding of their target population and directly sought to address those needs in their theory of change. In Camfed, for example, a holistic approach saw interlinked activities such as Learner Guides, *My Better World* materials and study guides complement one another in addressing specific barriers to girls' learning; while in BRAC, community-based girls' schools established by the project alongside stipends and support given to government school girls helped them begin to overcome specific barriers to their education.

- **Economic interventions, though effective in reducing the cost of schooling, are not enough to overcome the complex and multidimensional barriers of poverty.** Scholarships, bursaries and stipends are helping to reduce the cost of schooling while village loans and savings schemes, livelihood programs and cash transfers are all providing families access to additional income and financial support to send their girls to school. However, these interventions are not able to respond to households' basic needs or their ability to sustain the costs involved in sending their girls to school. Contextual factors, such as drought and insecurity, can also have detrimental impacts on livelihoods, the effect of which can mean families being forced to migrate, with girls having to drop out of school.

- **Upgrading sanitation facilities can be an effective way of encouraging more regular attendance among girls, as long as there is suitable access to water.** Projects seeking to improve infrastructure and sanitation facilities did not always consider the impact of water availability on the effectiveness of their interventions. And yet, according to data gathered by the EM, over half of treatment schools (57%) at endline were reported to be without access to water. A lack of water can mean that improvements in toilets and sanitation facilities are rendered obsolete if girls are not able to use them correctly. Girls can also be deterred from remaining in class if they are unable to access safe drinking water at school.

- **Interventions that support girls directly with their learning can improve their ability to perform well in literacy and numeracy.** Direct support with homework, reading and maths exercises can help girls grasp difficult concepts and improve confidence in their ability. Learning guides and books are also enabling girls to study at their own pace. Interventions such as child-centred teaching methods, reading clubs and study guides are all helping girls receive the direct support they need in order to learn effectively. Unless support is directly related to learning, however, it is unclear whether addressing girls' low aspirations and lack of self-confidence is linked to improved learning.

- **Mothers and community members can be effective champions of girls' education, helping girls to attend school.** Community members, mothers and caregivers are all playing an active role in mobilising support for girls' education and helping girls overcome individual barriers to being able to attend school regularly. Likewise, radios are proving effective in changing community perceptions and attitudes to girls' education, providing issues of access and radio signal have been addressed. However, despite their ability to mobilise support from others, evidence suggests that mothers and community members are not able to shift broader beliefs and behaviours around early marriage and pregnancy, or prevent girls from dropping out due to these factors. As identified in the GEC Thematic Research Report (2016), this may likely be due to the complex and multidimensional nature of poverty, early marriage, and livelihood support.

- **Confidence that violence issues will be reported is itself not enough to improve girls' sense of safety at school.** The threat of violence faced by girls is not something that SCW projects have successfully been able to overcome during the timeframe. While child protection policies and associated training have improved violence reporting mechanisms, EM data finds that one in four girls (25%) in treatment schools at endline reported feeling afraid at school. Ultimately, the effects of improving response to in-school violence may be too indirect to be perceived by girls themselves, at least in the short-term. These fears likely also reflect the persistent issues of harassment and violence faced by girls on their way to school, compounded by long distances and external security threats, which the projects have not been able to entirely overcome.

3.5 How scalable and sustainable are the activities funded by the SCW?

Key Findings

- Without the support of GEC-T it is highly unlikely that many SCW activities would have continued as they stand now and very few would have continued beyond the immediate short term.

- By the end of the GEC, projects were reliant on schools, teachers and community volunteers to continue implementing activities if further support, funding or remuneration was not available, for example, through GEC-T. While there is evidence of commitment to do this, in the longer-term this is unlikely to be possible, particularly for approaches and groups that are relatively new.

- There is some evidence of training undertaken by projects designed to facilitate the continuation of activities, for example through the use of cascade training, but this seems to have been carried out late and there is little evidence of the use of this training to develop or support new groups or learning.

- All projects have engaged with government at national, regional or local levels, with evidence of some projects benefiting from favourable policy environments.

- Some effective advocacy is evident and there has been some success in incorporating GEC materials into curricula and teacher training. These are important achievements, given that many projects' advocacy has been constrained by external factors or less supportive policy environments or because they started too late.

- Government buy-in to take over or run SCW project activities has not been achieved. Where projects planned for government to take over activities at the end of the project, for example in Afghanistan, there appears to still be confusion and questions over the transition process and the allocation resources to enable this to happen.

Approach to assessing sustainability in the SCW

The business case recognises that for all windows (not just the SCW) there was a risk that there would be a trade-off between sustainability and achieving short-term results. The business case states *"achieving sustainability (systemic change) from interventions is very hard to do, and so there is a risk that funds naturally flow to organisations who are able to clearly demonstrate short-term results but are less able to articulate the long-term sustainability of interventions.*" and notes that "*in practice, there will need to be an understanding of how funding will be sustained beyond the programme period, which will require some implicit commitment to sustain support until alternative sources are in place.*"[103] This suggests that DFID considered that sustainability would only be achieved if the programme was extended beyond the programme period.

From the outset SCW projects were expected to identify and report on the mechanisms that they would put in place to sustain their activities at the end of the GEC. The GEC programme finished in March 2017. GEC Transitions (GEC-T) has been developed by DFID and the current FM as a successor to the GEC. It will provide further funding and support for those GEC projects that are able to meet and deliver GEC-T's objectives, with a focus on continuing to provide support for the cohort of girls who benefited from the first phase of the GEC. As such, GEC-T potentially provides SCW projects with more funding (and time) to enable them to find alternative resources to continue and sustain the activities and benefits delivered during both the GEC and GEC-T programmes. As the EM for the GEC we have not been involved in the development of GEC-T to date, which means that we are not able to comment further on the potential sustainability of GEC projects through this successor programme.

This section relies exclusively on the projects' endline reports. In their endline reports projects were asked to respond to the following questions:

- Has the project put in place mechanisms that allow changes to marginalised girls' attendance and learning to be sustained?

- How likely is it that the projects' benefits will be sustained?

- To what extent has the project leveraged additional investment to sustain its activities?

The majority of projects followed this guidance in their endline reports and all reports include information on sustainability. However, the quality and depth of reporting varied across the reports, and where required we have also drawn on material from the report effectiveness sections. It is important to note that not all projects' activities are reported in the sustainability sections of the endline reports, so there may be further insights from the projects themselves. Overall, there was a lack of reported evidence of direct feedback from project staff, beneficiaries and stakeholders in projects' endline evaluation reports. We have taken the information provided and highlighted the most useful insights into the opportunities and challenges that the projects face with regards to sustainability, as well as the typical characteristics of the SCW projects' sustainability.

Annex L provides summary tables of the evidence about sustainability that we extracted from each of the projects' endline evaluation reports. We used the following criteria to extract and analyse evidence of projects' sustainability at endline:

- Information about which **type of activities proved most effective** and as a result should be sustained;

- The cash or in-kind **resources that the project has leveraged** specifically to sustain activities after the end of the project;

- Information about **who would be responsible** for sustaining key activities;

- Evidence of **key drivers** sustaining the project's activities;

- Evidence of **key barriers** to sustainability;

- A **judgement** about how likely it was that each of the activities identified would be sustained given the evidence available; and

- A **judgement** about the sustainability of the project as a whole.

We also categorised different parts of projects' sustainability strategies at global, national, regional, school and community levels. Our judgements of the projects' overall sustainability are presented using a RAG rating system. As indicated in

Table 18, a green rating indicates strong performance, a fully sustainable project with strong support and funding in place, whereas a red rating indicated that no viable evidence of sustainability for most of the activities was presented in the endline reports. The RAG ratings are not particularly scientific - rather, they allowed us to make reasonably evidence-based judgements about the potential sustainability of each project as it stands now, for the purpose of informing a general finding about the sustainability of the SCW portfolio as a whole. The ratings developed for each individual project are provided in Annex L. Our overall assessments of projects' sustainability is broadly in line[104] with the assessments made by the FM in the GEC Project Completion Report, which emphasises the need for a more structured approach to sustainability for GEC-T.

---

[104] We have used a RAG rating approach to enable us to systematically assess the sustainability of the SCW portfolio as a whole. The FM has not rated the sustainability of each of the SCW projects in the same way.

Table 18: RAG rating system to assess sustainability

Rating Symbol	Rating description
G	**Green:** strong performance, evidence of a fully sustainable project with support and funding in place to continue all desired activities
G A	**Green – Amber**: performed well, evidence of support and/or funding in place to support major or a majority of activities. Or evidence of beneficiaries independently running activities
A R	**Amber – Red:** underperformance. evidence of work towards sustainable activities but major or a majority of activities are not sustainable
R	**Red:** poor performance. No viable evidence of sustainability for any project activities

A key precondition to sustainability is knowing which activities or interventions are the most effective in delivering an impact for marginalised girls (and as a result are the most critical to sustaining the delivery of long-term benefits). Just one project, Plan (Sierra Leone), included a discussion around the recognition that sustainability of the project's impact would involve scaling up or scaling down different aspects of the project.

We found limited evidence of additional financial resources being obtained to sustain project activities after the GEC, excluding GEC-T. The sole example was funding from UNHCR secured by WUSC to continue to provide the scholarships. Otherwise, projects are largely reliant on officials, teachers, and community members continuing to deliver activities, or adapted activities, without support or remuneration.

Key findings across the SCW

All the projects reported that they had engaged with **governments at national, regional and district levels** to support the delivery of activities, with some success in terms of aligning with policy and influencing change. As discussed in more depth in Section 3.4, projects in Zimbabwe, Kenya and Tanzania seem to have benefited particularly from a positive policy environment with positive potential impacts for sustainability where projects were able to comprehensively train officials (WV). However, this engagement has translated into few concrete plans for ongoing funding of activities or government takeover of activities. It is important to note that the full incorporation of project activities into national systems was not realistic within the timeframe of the GEC.

At midline, projects seemed to be at an early stage in planning and implementing their sustainability strategies. It was generally not clear what their strategies were for influencing government behaviour and with what specific effects on the sustainability of the project activities. Many projects spoke of refocusing on advocacy for governments to take over parts of their project in its final year. At endline there is evidence of projects' engagement with government through advocacy, participation and leading working groups or committees, and the production of advocacy materials, but there was little evidence of this leading to concrete financial support from governments for the continuation of project activities. The most substantial influence on policy appears to be: Save the Children's (Mozambique) influence on government who is planning to fund the Literacy Boost programme from the national budget; and CARE (Somalia), whose recommendations for the inclusion of a gender perspective in school inspection checklists and the provision of training to officials and head teachers are being adopted by government. However, CfBT (Kenya) also reported that the TUSOME programme had raised questions of duplication and waste of resources, prompting the project to pull out of targeting early grades, following a decision made by the Ministry of Education.

The largest commitment to take on GEC activities seems to have been secured by the GEC projects in Afghanistan (AKF, ACTED, and BRAC). The sustainability strategies for these projects were based on transitioning activities over to the government. However, while it seems that there is a commitment from the government to continue to pay teacher salaries for the ACTED project, more comprehensive takeover plans seem to remain uncertain. For instance, BRAC's report mentioned that officials showed limited awareness of handover plans and raised concerns about the availability of resources to enable children to attend government schools. According to their report, no funds would be available to establish outreach classes as had been previously committed. Ongoing challenges with security on the journey to school also increase the barriers to girls' transitioning from community to government schools. No additional resources seem available to provide transport or to support community schools closer to girls' homes.

> "*Yes, few numbers of girls will go to government school after finishing the community school because we have insecurity problem in our country for example; in the previous month some of the government school girls have been kidnapped and after three months we just found their corpses.*"

> DED Official, Jabul Saraj District, Parwan (BRAC Afghanistan)

Projects have engaged with local and district officials to build awareness and commitment to their activities. Sustainability strategies focused on capacity building and involved district and local level officials in monitoring project activities. There was strong evidence presented that officials generally had positive attitudes towards projects and supported activities (CARE (Somalia), WV (Zimbabwe)). Other projects increased the capacity of local officials in M&E with the hope that they will continue and expand these activities. For example, ChildHope (Ethiopia) planned that woreda officials[105] would take on the administration of the project's M&E framework for GEC-T. However, there is no clear evidence that an agreement had been reached in their endline evaluation report.

Projects are very reliant on schools to continue to support activities, maintain facilities, and ensure that improved teaching and management practices are maintained. They are **reliant on teachers** to continue using the new techniques they have learnt. But projects do not always take into account the impact of teachers' transfers and the lack of financial support and ongoing training.

Most projects included teacher training as discussed under *Teachers' inadequate pedagogy and teaching skills*, covered in Section 3.4. Teachers seem to generally be supportive of the changes and have adopted new teaching practices, although there are concerns that they are not always able to recall the specific training that they have received (CARE Somalia), with ongoing training and support likely to be required. In addition, schools have adopted improved child protection policies, and there is some evidence that children feel more confident in reporting abuse directly to teachers (ChildHope (Ethiopia)). Projects have worked to establish and build the capacity of school management committees to support and drive these changes, but there are few plans for ongoing support or training of new committee members to ensure that the skills and focus remain. Save the Children (Mozambique) is an example of a project that has used a cascade model to train their School Safety Committees: training materials are available to committee members which gives potential for the ongoing transmission of these skills.

Teachers often comment that they will continue to use new methodologies after the completion of the project. However, in most cases there do not seem to be any concrete plans to: ensure the ongoing use of new methodologies; deal with high levels of teacher transfers; ensure that more teachers in the school are trained; or ensure that in some cases salaries will continue to be paid (BRAC and ACTED (Afghanistan), WUSC (Kenya)). Projects' sustainability plans also seem to be overly reliant on the production of manuals and materials. Some projects are passing these on to governments and other NGOs with the expectation that these will serve as sustainable learning about good practice (IRC (DRC), AKF (Afghanistan)). Camfed (Zimbabwe and Tanzania) claimed that the Learner Study Guides and Life Skills Training curriculum that had been developed have been rolled out across four further districts. Camfed (Zimbabwe and Tanzania) also suggested that study guides were being incorporated into wider classroom delivery, with the effect of encouraging teachers to employ different

---

[105] Woredas are local level administrative regions in Ethiopia, similar to a district.

methods of teaching and make greater use of child-centred pedagogies. This is promising evidence that synergies between interventions may help increase the likelihood of delivering sustainable outcomes.

Projects are also broadly reliant on teachers or other volunteers to continue to run girls' clubs and extra curricula groups without remuneration or ongoing support. This is likely to mean that the continuation of activities is short lived. Plan (Sierra Leone) trained an additional cohort of Programme Volunteers to run new study groups at 180 schools. There are some reports of new study groups forming independently of the project, but there is no evidence of securing the financial support that they require to run these groups or provide materials. Camfed (Zimbabwe and Tanzania) also indicate that the Learner Guides had established study groups outside of school hours, which was not envisaged in the design of the project. While they welcomed this, they also cautioned against the risks of over-burdening girls with school administration requirements. The most sustainable girls' clubs in the SCW projects are likely to be those run by CfBT (Kenya), which have strengthened existing clubs rather than created new ones. Group members and leaders have expressed a commitment to continue activities.

There is very limited evidence of sustainability plans for material, in-kind and financial support or to maintain infrastructure established during the GEC. WUSC (Kenya) provided evidence that some funding has been raised to continue scholarships, but no other donor or NGO contributions are evident. Community fundraising may be more sustainable in the long term where community incomes make this viable, but there are concerns over the rising cost of tuition fees relative to income, which makes the question of sustainability of these interventions more challenging.

IRC (DRC), AKF and BRAC (Afghanistan) all provide evidence of school management committees or communities raising funds for the construction of latrines, or to repair school buildings, while WUSC (Kenya) mentioned that their sanitary wear model had been adopted by other international NGOs operating in the area. However, CfBT suggested that no uptake mechanism was in place to continue with the provision of uniforms (as part of back-to-school kits) and both WUSC (Kenya) and RI (Somalia) raised issues around the upkeep of solar lamps that had not been considered in their project design. The general sense that activities are not likely to continue without ongoing financial support is summed up by this community leader:

> "*KEEP was to end, enrolment, performance would all go down because there will be no uniforms, textbooks, solar lamps for night studies. There will be dropouts because KEEP is the one providing scholarships, sanitary wear, pays teachers. If KEEP pulls out, many girls will stay home. Absolutely no resources for maintenance of these inputs in this community unless provided by a project like KEEP.*"
>
> Community Leader, WUSC (Kenya)

Training activities for unqualified teachers have had a strong positive impact that could have long-lasting effects. In the case of ACTED (Afghanistan), it increases the likelihood that literacy education will continue to be provided at the community level without financial support from the government. Camfed (Zimbabwe and Tanzania) trained 3653 young women to become Learner Guides across the two countries. Many of these guides have secured places at teacher training colleges and have been able to improve their academic qualifications and become engaged with community decision-making structures as a result of the confidence gained from being involved in the programme.

Projects also need to be able to cope effectively with challenges like the increased demand for education because of awareness-raising activities. ACTED (Afghanistan) report overcrowding in classroom and increased violence, which potentially jeopardises continued support for education. Besides this, problems with the quality of some of the buildings were already evident at endline.

Communities are generally supportive of the GEC and there is evidence of positive changes in attitudes towards girls' education. However, there has been **little planning for how community activities will continue**, or capacity building to enable communities to run activities themselves. Poverty reduction activities do not appear to have been successful enough generally to guarantee that parents will continue to spend any additional income on school-related expenses.

As discussed in more detail in Section 3.4, there is evidence that community support for girls' education has improved. WV (Zimbabwe) reported evidence that proactive mothers' groups were effectively supporting girls'

education and helping to prevent early marriage. While respondents who have been involved in community activities often stated that they would continue to be involved (WV (Zimbabwe)), or as in the case of Save the Children (Mozambique) have been provided with some training to deliver advocacy messages and radio broadcasts, these activities are unlikely however to continue without continued support and awareness-raising. Reports did not describe or explain what formal or informal infrastructure is available to continue supporting voluntary groups to sustain projects' activities. Critical mass and attitudinal change can certainly lead to the behaviour change required to sustain new types of educational support activities, but there is little evidence that this has been achieved by endline. Another concern is sustaining the engagement of community members who feel that they have already changed their views, as was reported to be the case in BRAC (Afghanistan), where participation in community workshops waned over time as community members no longer felt the need to attend the meetings once they perceived to have benefited from the relevant information.

Similarly, while community members have actively engaged in poverty-reduction activities like VSLA groups, there is little explanation about how these groups are expected to continue to function without ongoing support, particularly where there have been challenges in successfully running businesses. Examples include:

- IRC (DRC): EA$E VSLA are successful, with members repaying loans and using the income to pay school fees. Community members have been trained to run the activities but community leaders express concern about whether this is sustainable in the long term without additional support, and whether incomes will continue to be high enough to pay school fees;

- WV (Zimbabwe): 91% of the groups are meeting regularly and 61% engaged in income-generating activities. Community members have been trained to run the groups but there is a perception in the communities that groups are being wound up; and

- Child Hope (Ethiopia): self-help saving groups have been successfully transferred to be run as government supported cooperatives.

Some SCW projects will continue to receive support through the successor programme to the GEC, the GEC Transitions (GEC-T) programme. Without this support, it is highly unlikely that many SCW activities would have continued as they stand now, and very few would have continued beyond the immediate short term. The extent to which girls, schools and communities will continue to benefit from GEC-T will depend on the subsequent design of the projects which are now tasked with continuing to support the same beneficiaries as they did in the GEC.

3.6 To what extent does the SCW represent good value for money?

Key Findings

- SCW projects prioritised their investments in interventions that address the most pressing needs facing girls' education i.e. poverty and school-related barriers.

- Expenditure on *economic* interventions addressing poverty-related barriers are associated with delivering the most learning beneficiaries. This is consistent with our findings in Section 3.4 that in-kind support improved attendance and bursaries, loans and savings and income-generating activities helped improve retention, attendance, and indirectly, learning.

- Expenditure on school-related interventions appears to have delivered the next largest number of learning beneficiaries; that is, *teacher training and support*, followed by expenditure on *infrastructure and resources* and then *school management and governance*, which is also consistent with our qualitative analysis of what works.

- Amongst all poverty and school-related interventions, *teacher training and support* and *school management and governance* interventions are the most cost-effective at on average £64.70 (*teacher training and support*) and £56.20 (*school management and governance*) per learning beneficiary.[106] *Economic* interventions addressing poverty-related barriers and interventions improving *infrastructure and access to resources* are the most expensive out of the SCW interventions, at £123.90 and £104.60 per learning beneficiary, respectively.

- Amongst all interventions, *extra-curricular activities and non-formal educ*ation seem to be the most cost-effective for girls' learning. This finding is supported by effectiveness evidence that activities aimed at improving learning directly, such as through tutoring clubs or accelerated learning programmes, have had a larger positive effect on girls' learning than 'indirect' activities such as scholarships or discussing the benefits of staying in school.

- The average unit cost per beneficiary for numeracy is about two to three times the average unit cost for literacy.

- At the end of the GEC, SCW interventions have not delivered the amount of improvement in literacy and numeracy to ensure the marginalised girls progress through their schooling at the right rate and with grade-appropriate levels of learning, despite demonstrating projects reaching and benefiting a large number of girls.

- More data is needed about the characteristics of subgroups to inform VfM assessments that help guide projects and the programme about what works, for whom, under what conditions and at what cost.

- While projects provide activity-based costing that is associated with different types of interventions, more explicit linking of interventions to outcomes by projects is needed to help assess more accurately which types of interventions deliver the best value for money. It is not necessarily the case that the magnitude of learning impacts is directly proportional to the magnitude of expenditure on certain interventions. We had to work under the assumption that the quality of the design and delivery of different types of interventions were the same within and across projects, which is not the case.

Approach to assessing value for money

Our value for money (VfM) assessment of the SCW at midline was solely based on the VfM assessments included in projects' endline evaluation reports. These were generally of very poor quality. The FM, EM and DFID therefore agreed that projects' endline evaluations would not need to include a VfM assessment. Instead, the EM would use the available impact data (reported by projects) and programme finance data (collected from projects and reported by the FM) to conduct a VfM assessment at the window level.

Our approach to this VfM assessment is as follows:

---

[106] Unit cost is calculated by dividing total expenditure per intervention by each project's learning (average of literacy and numeracy) impact relative to 0.5 Standard Deviations – each project's baseline to endline target – and then dividing by the total number of beneficiaries reached.

- **We focus on VfM at the window level.** This means that we are looking at different types of interventions, their relative effects and costs. As far as possible, we have not undertaken cost-effectiveness analysis for the purpose of comparing individual projects and drawing VfM conclusions to identify which projects delivered best value for money. The differences in: their contexts; the types of marginalisation they are tackling and the type of girls they are targeting; the type of interventions they are delivering; and their scale makes it difficult to draw accurate and meaningful VfM comparisons between projects. This is especially the case when using cost-effectiveness analysis. We did however, explore VfM at the project level and for reference purposes only we have included our analyses in Annex M.

- **We have used cost-effectiveness analysis to assess VfM at the window level** by comparing the efficiency of different ways (interventions) in which resources have been used across the SCW to impact on the literacy and numeracy outcomes of the marginalised girls targeted by projects. We could not use cost benefit analysis because of a lack of relevant and accurate data to measure the lifetime social and economic returns on different types of investments in marginalised girls' education across diverse contexts.

- We assess VfM by:

    - reviewing the **scale of investments** across different types of interventions;

    - assessing the **scale of the impact** in terms of the number of girls benefiting from learning gains;

    - assessing the **cost-effectiveness of different types of interventions** through analysis of the average unit cost of improving the literacy and numeracy of one girl by 0.5 Standard Deviations over and above the control; and

    - using these assessments together with key findings about the effectiveness and sustainability of the SCW to **draw broad conclusions** about the VfM of DFID's investment in the SCW.

- **We recognise that this approach is crude and imperfect** and as such the analysis and findings should be interpreted with caution. The key methodological limitations are as follows:

    - *Data was missing and /or incomplete for some projects*, most notably for ChildHope (Ethiopia), Relief International (Somalia), World Vision (Zimbabwe), IRC (DRC), ACTED (Afghanistan) and Plan (Sierra Leone). There were also some disparities between the total actual expenditure reported by the FM and the total expenditure reported in the VfM cost data from projects (see Table 55 in Annex M).

    - *Standardising the degree of girls' marginalisation* – in this assessment we assume that the level of effort necessary to increase the learning of girls is the same across all projects' contexts. This is, of course, not true. Girls are subject to various types of factors that marginalise them from education in different ways and to different degrees of severity. A standard level of improvement in learning for one girl will most likely require a different level of effort (and cost) compared to the same level of improvement for another girl in another context. This affects the type and scale of costs that are necessary to achieve the same level of improvement for each individual girl or group of girls.

      This is a key reason why it is important to collect *data about the characteristics of different subgroups* to help us evidence and understand what works, for whom, under what conditions and at what cost. There is currently an insufficient amount of data about subgroups to systematically conduct analysis of the distribution of different effects on girls who are marginalised in different ways and to different extents.

    - *Not all learning tests are comparable* – we have to rely on projects' impact achievements which, as shown in Section 3.3.5, are not necessarily comparable across projects who used different methodologies (different tests, different evaluation designs).

    - *Assumption that cost categories are mutually exclusive* – the same types of activities do not necessarily fit solely into one type of intervention or the other. Moreover, projects have holistic and often complex approaches to their theory of change (ToC). This may distort the relative scale of investments in different types of investments at the window and project levels.

The main difficulty in conducting the cost side of the VfM analysis was classifying different expenditure activities. Inevitably, activities could only fall into one category, which was a subjective judgment on our part (although we did harvest this information from projects' endline evaluation reports). This might distort the projects' intended ToC. For instance, we decided to associate *community-based* interventions with attitudes to align with the way we structured our evaluation of the SCW's effectiveness.

- *Scale of investment does not necessarily correlate to an effective intervention* – it is possible that relatively small investments in some types of interventions may have a large effect or that some projects had competitive procurement practices and operational strategies in place to achieve the best price for inputs. While there was no reporting from projects on their cost decision-making and we did not find explicit evidence of differences in effect sizes between expenditure across interventions, care is needed in our assumption for the purpose of this assessment that the scale of investment correlates with the scale of effect /benefit.

- *Wider 'value' (effects and benefits) is not captured* – value is defined in this assessment as the reported impacts on girls' literacy and numeracy. Unlike cost-benefit analysis, cost-effectiveness analysis is not the right tool for aggregating and comparing all lifetime costs with all lifetime benefits. For practical reasons, projects have not systematically quantified indirect benefits, spillover effects or multiplier effects, or conversely displacement and substitution effects. Furthermore, intangible (qualitative) benefits are also not captured through this type of VfM assessment.

- *Future benefits that may be realised with more time are not captured* – some types of interventions may take longer to result in an observable effect on learning than others.

- *The persistence effects of different types of interventions are not captured* – this relates to the extent to which a girl will continue to benefit after benefiting from the initial intervention. To some extent this may be affected by whether or not activities are likely to continue after the end of the project. But it is also affected by the type of capacity a girl has developed as a result of an intervention and the extent to which this enables her to progress at the same trajectory through her schooling and beyond into adulthood.

Despite these limitations, we believe that this assessment provides a useful overview of the VfM of the SCW as a whole and complements the findings presented throughout the rest of the report.

SCW expenditure by type of intervention

Figure 19 shows the absolute and relative investments made in different types of interventions across the SCW. It is useful grouping these types of interventions by the barriers that have framed our evaluation of the effectiveness of the SCW, to enable us to compare our VfM analysis with what projects reported worked well or less well. It is important to note (as above) that in categorising different types of activity costs we have had to make subjective judgements, which may distort the actual picture of what projects invested in and delivered on the ground. However, given that caveat, the SCW has invested:

1. 46% (£89.5m) on *school-related* barriers, consisting of:
   a. 19% (£37.3m) on *teacher training and support*.
   b. 17% (£33.3m) on *infrastructure and resources*.
   c. 10% (£19m) on *school management and governance*.

2. 25% (£47.4m) on *poverty and livelihoods* barriers, consisting of *economic* types of interventions.

3. 8% (£15.9m) on *attitudinal barriers towards girls' education*, consisting of *community-based* types of interventions.

4. 8% (£15.3m) on barriers relating to *girls' aspirations*, consisting of:
   a. 6% (£12.5m) on *extra-curricular and non-formal education*.
   b. 1% (£2.8m) on *empowerment and self-esteem* interventions.

5. 0.4% (£0.7m) on barriers relating to *violence and safety.*

Figure 19: SCW Absolute and Relative Expenditure by Type of Intervention

All SCW projects*
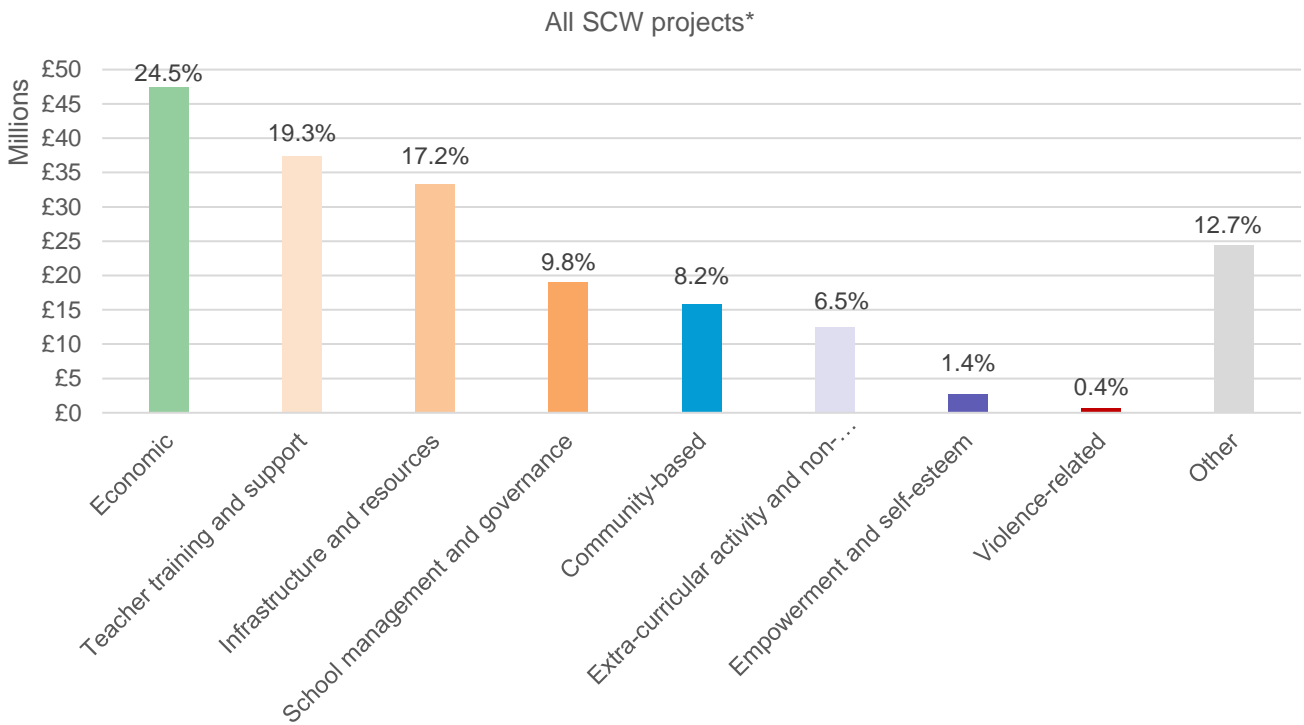
* No data from Plan (Sierra Leone)
† Other = operation and overhead expenses, HR, salaries & admin costs, M&E activities, 'other'

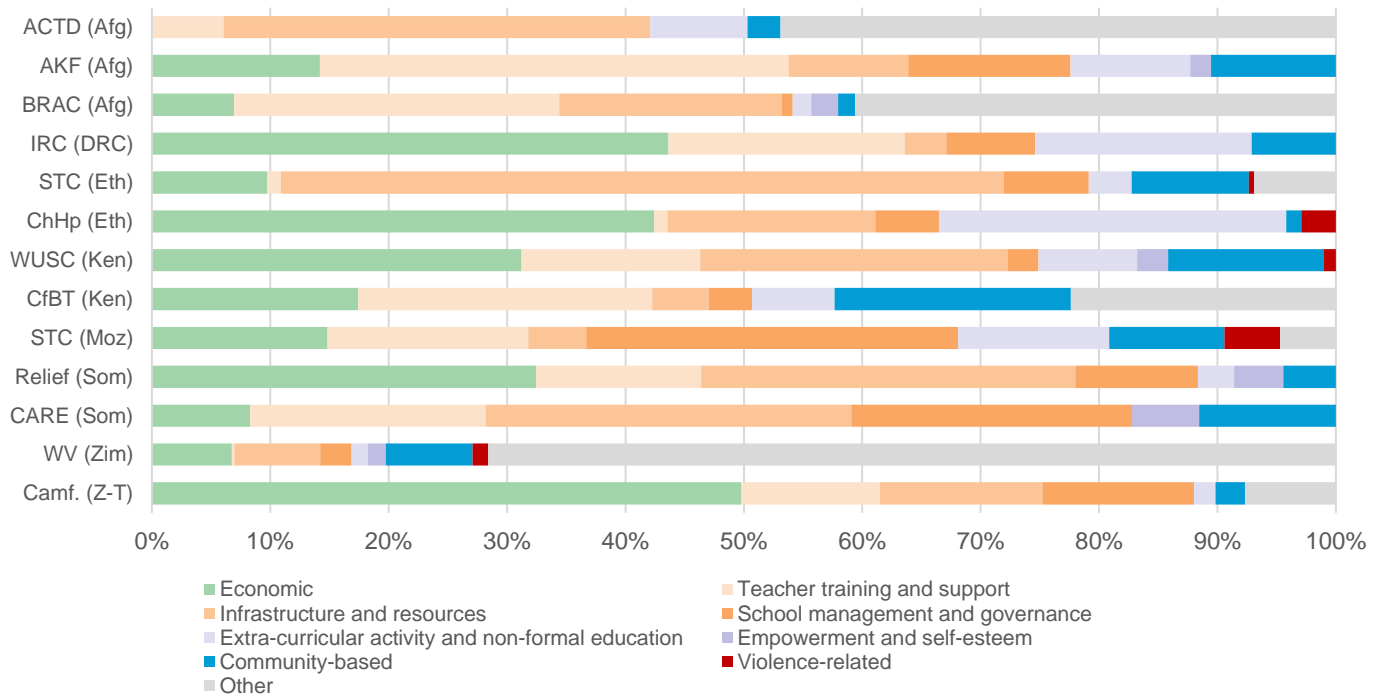## Figure 20: SCW Relative Project Expenditure by Type of Intervention

Figure 20 shows the individual investments made by each of the projects by intervention type. Thirteen projects provided VfM metrics data. Plan (Sierra Leone) did not provide any data and it should be noted that IRC (DRC) and ACTED (Afghanistan) VfM data only covered the period up to their midline evaluation report. Project expenditure

data simply confirms the pattern of expenditure at the window level showing that roughly two-thirds of all projects' expenditure was on interventions targeting poverty and school-related barriers. Our qualitative analysis shows that these two barriers remain (since baseline and midline) the two most important barriers to marginalised girls' education. This data also confirms that SCW projects prioritised their investments in interventions that address the most pressing and evident needs facing girls' education.

It is worth recapping here the key findings from our evaluation of what worked (Section 3.4) based on the evidence submitted by projects in their endline evaluation reports:

- *Economic interventions* and in-kind support helped to reduce the cost of schooling and led to improved attendance among girls.

- *Infrastructure and facilities* improvements have positively impacted on girls' attendance, while learning materials led to improvements in girls' learning.

- While projects presented evidence of improvements in the *quality of teaching and teaching methodologies* at endline, only one project implementing teacher training and support interventions presented conclusive data explicitly linking these to improvements in girls' literacy and numeracy (ChildHope (Ethiopia)).

- Very little mention was made of the effects of *school management and governance activities* on attendance or learning outcomes. It was noted that there has been an increase in psychological support provided to girls at school, but this was not linked to changes in educational outcomes as a result.

- *Community-based activities* such as gatherings and meetings have encouraged caregivers and community members to take an active role in supporting girls' education, leading to improvements in girls' attendance.

- *Extra-curricular activities* have raised girls' self-esteem and confidence in their academic ability, with some improvements in learning outcomes from activities that directly targeted girls' learning, such as tutoring and accelerated learning programmes.

- Interventions addressing *violence* have not had conclusive effects on girls' attendance or learning outcomes.

Interventions designed to address school-related and poverty barriers appear to have been most effective in improving literacy and numeracy on the basis of our qualitative analysis of what has worked well, as reported by the projects. Projects have spent most of their budgets on these types of interventions – 46% of all spending was on school-related interventions and 25% on poverty-related activities. In analysing cost-effectiveness, however, it is important to also consider the number of girls reached and benefiting from different types of SCW interventions, as cost is typically correlated with scale, i.e. it broadly costs more to reach and benefit more girls.

The number of girls benefiting from literacy and numeracy gains in the SCW

Figure 21 shows the total number of *learning* beneficiaries reported by the FM as an average of the baseline to endline impact on literacy and numeracy (percentage achievement against target) multiplied by the number of girls reached by each type of intervention.

When grouped together, expenditure on school-related interventions appears to have delivered the greatest number of learning beneficiaries. Notably, *teacher training and support* activities are associated with a substantial number of learning beneficiaries. This contradicts the lack of evidence about the effectiveness of improved teaching methodologies specifically in improving literacy and numeracy, although it is possible that this is the result of a reporting deficit (in that no evidence does not necessarily mean no impact). Given the sheer amount of resources allocated to *teacher training and support*, it is unlikely that there would be no effect on a reasonable share of girls' educational outcomes.

Expenditure on *economic* interventions addressing poverty-related barriers are associated with delivering the most learning beneficiaries by individual intervention. This is consistent with our findings in Section 3.4 that in-kind support improved attendance and bursaries, loans and savings and income-generating activities helped improve retention, attendance, and indirectly, learning. By offsetting the cost of education, economic interventions helped increase attendance (as an intermediary outcome) and facilitated learning by enabling girls to engage in an academic environment.

Expenditure on *infrastructure and resources* interventions appears to have delivered the third largest number of learning beneficiaries. This is consistent with our qualitative analysis that improved infrastructure and sanitation facilities encouraged attendance, and learning materials directly improved girls' literacy and numeracy.

Examining interventions by the scale of their impact, the three most expensive interventions brought in by far the greatest number of learning beneficiaries. Yet while it is useful to examine which interventions reached the most beneficiaries, in analysing value for money it is also important to combine cost and scale analyses to determine which interventions resulted in the most learning gains at the least cost.

Figure 21: Number of Learning Beneficiaries per type of intervention[107]

* Extra-curricular activity and non-formal education
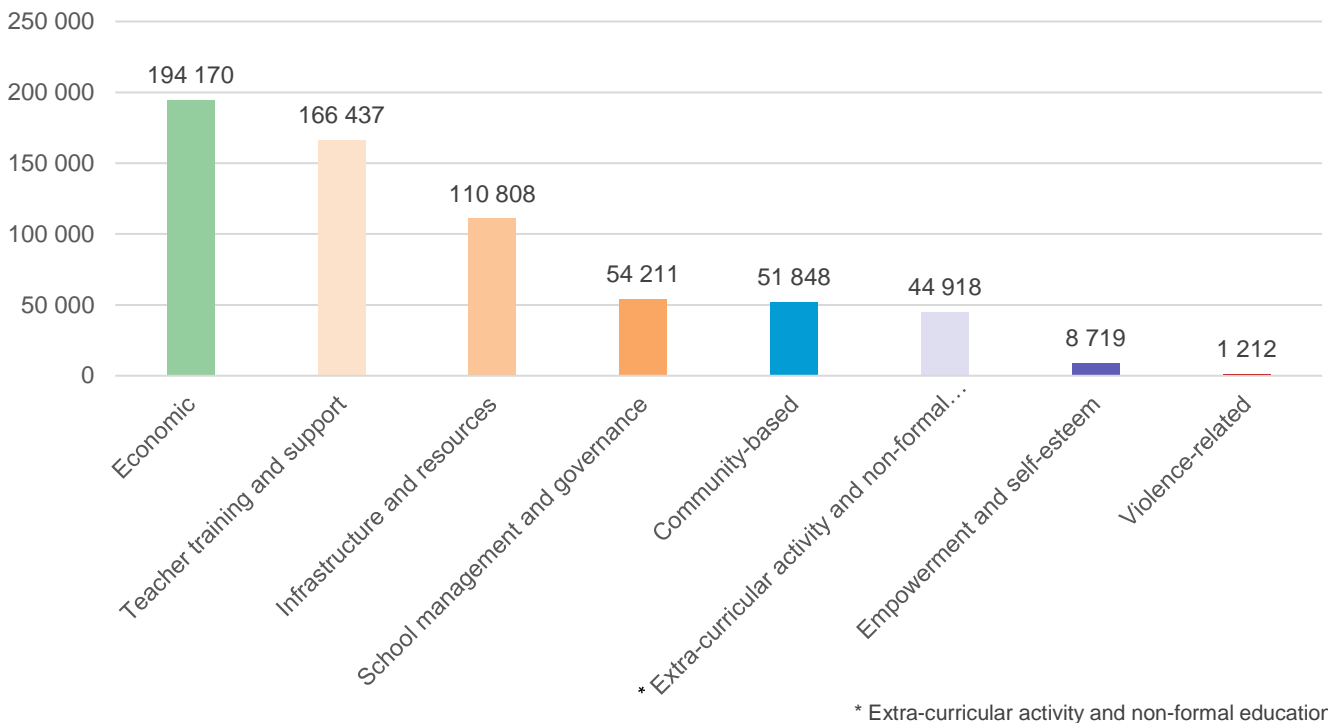
Cost-effectiveness of the SCW interventions

Figure 22 and Figure 23 show the average unit cost per literacy and numeracy beneficiary, respectively, relative to expenditure on each type of intervention. This was calculated by dividing per project the total cost of each intervention, adjusted for overhead costs[108], by the number of learning beneficiaries – calculated as the number of girls reached multiplied by twice[109] the impact achieved in standard deviations – and then averaging across the window.[110]

The ranking of the most and least cost-effective interventions is slightly different between literacy and numeracy. While the most cost-effective intervention for both is *extra-curricular activities and non-formal education*, for literacy *teacher training and support* had an identical unit cost, whereas for numeracy the unit cost for *teacher training* was more than double that of *extra-curricular activities*. The implications of this finding are twofold. Firstly, that activities aimed at improving learning directly, such as through tutoring clubs or accelerated learning programmes, have had a larger positive effect on girls' learning than 'indirect' activities such as scholarships or discussing the benefits of staying in school. Secondly, this finding is consistent with the evidence of what has proved effective, and education literature suggesting that it is more difficult to teach numeracy than literacy. Across the window, SCW projects have

---

[107] 'Other', which captured operational, overhead and M&E costs, was redistributed among the remaining interventions based on their share of total expenditure.
[108] See previous footnote.
[109] This was a means of calculating the cost of a 0.5 Standard Deviation (SD) improvement for a single girl. We use 0.5 SD because this was equal to the most commonly used baseline to endline (lifetime) targets that were set for SCW projects.
[110] Our value for money analyses excludes STC (Ethiopia), whose impact data is inconclusive, and RI (Somalia), who had a negative impact in both literacy and numeracy. These projects should in theory have 'infinite' unit costs, and were therefore omitted so as not to skew our calculations. Given this, average unit costs across SCW are understated, both at the aggregate and unit cost levels.

been more cost-effective in achieving improvements in literacy than numeracy, with unit costs per intervention for numeracy being on average two to three times those of corresponding interventions for literacy.

In determining cost-effectiveness, it should be noted that while cost-effectiveness calculations suggest that *violence* and *empowerment and self-esteem* interventions are the most cost-effective, the scale of their impact in terms of the number of girls benefiting from improved learning was relatively small. *Violence-related* interventions accounted for less than 0.5% of total expenditure and only 0.2% of all learning beneficiaries across the window, while *empowerment* activities constituted 1.0% of spending and 1.4% of learning beneficiaries. Thus, while these interventions may be strictly speaking the most cost-effective, for the purposes of bringing about large-scale improvements in literacy and numeracy it is unlikely that these activities would proportionally scale to higher magnitudes of learning beneficiaries as a result of greater investment in these types of interventions.

Economic and *infrastructure and resource* interventions stand out as the least cost-effective types of activities. This is consistent with our analysis of projects' evidence about what worked. *Economic* interventions, such as bursaries and scholarships support individual girls, but they cost more on average to support each girl compared to interventions that are designed to benefit larger groups of girls. While improvements in infrastructure can affect whole schools, this is an indirect intervention primarily facilitating attendance, which is reliant on other factors such as the availability of good quality teaching, teaching materials, safe water supply, etc. to improve girls' learning. In terms of improving learning within the GEC three-year timescales, providing *infrastructure and resources* and *economic* interventions to address poverty is relatively less cost-effective compared to targeted interventions that are designed to more directly improve girls' learning. Of course, it should be noted that this finding relates to the learning results achieved within a relatively short three-year timescale. Interventions improving access to *infrastructure and resources* and *economic* activities may have effects on learning gains that take longer to materialise and as such are not as apparent within this programme period.

While acknowledging all of the caveats around the limitations of this assessment, the evidence suggests that at the window level *extra-curricular activities and non-formal education*, in addition to *teacher training and support* for literacy, are the most cost-effective interventions after taking into account the number of girls benefiting from learning and the effect size that has been achieved – their respective unit costs are £47.10 to £47.80 for a 0.5 Standard Deviation improvement in literacy and numeracy achieved per girl.

Figure 22: Literacy Cost Effectiveness: Average unit cost (£) to improve the literacy of one girl by 0.5 SD over and above control – excluding STC (Ethiopia) and RI (Somalia)
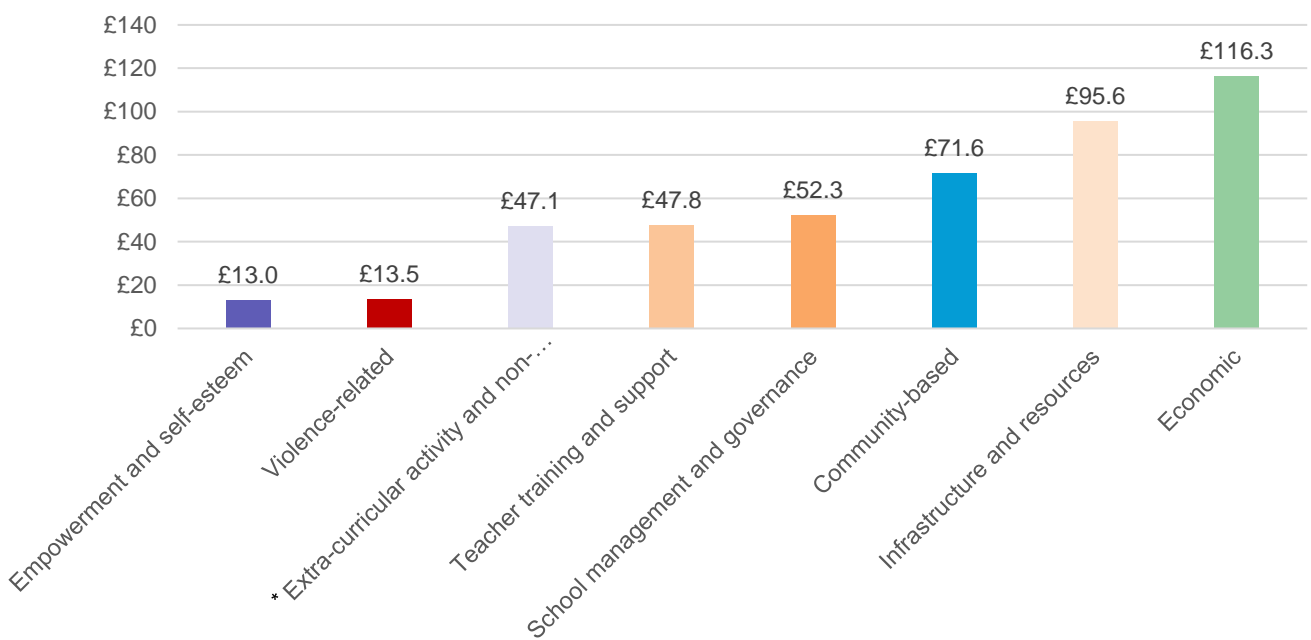
* Extra-curricular activity and non-formal education

Figure 23: Numeracy Cost Effectiveness: Average unit cost (£) to improve the numeracy of one girl by 0.5 SD over and above control – excluding STC (Ethiopia) and RI (Somalia)
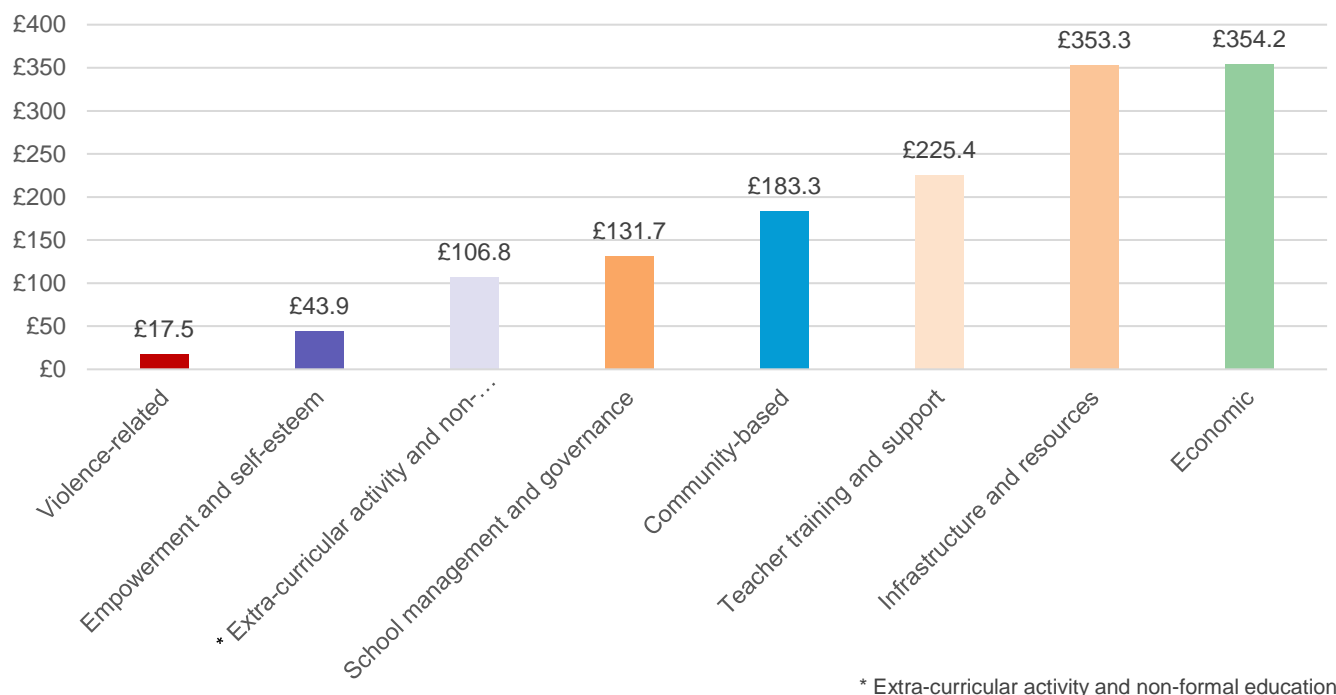
* Extra-curricular activity and non-formal education

Conclusions about the VfM of the SCW

This VfM assessment is based on analysis of what projects across the SCW have spent their budgets on, what effect this has had on improvements in literacy and numeracy, and the magnitude of impact on the number of marginalised girls benefiting from expenditure on different types of interventions.

School-related interventions and poverty-related interventions appear to have benefited the greatest number of girls. Our analysis of sustainability across the SCW suggests there are significant issues, though, in sustaining these activities (in particular *economic* interventions), which potentially constrain the long-term or lifetime benefits that could be realised from these interventions. Furthermore, in terms of their cost-effectiveness, *economic* interventions addressing poverty-related barriers and interventions improving *infrastructure and access to resources* are relatively more expensive. The evidence suggests that targeted interventions providing *extra-curricular activities and non-formal education*, as well as *teacher training and support* for literacy outcomes, are the most cost-effective. It is difficult to assess how cost-effective these results are in the absence of other unit costs for similar types of interventions delivered by other projects /programmes.

At the end of the GEC, our analysis of the impact of SCW projects suggests that these interventions have not delivered the amount of improvement in literacy and numeracy to make them more cost-effective, despite demonstrating some effects that have benefited the number of marginalised girls that were targeted at the start. It may be that more time is needed for these interventions to deliver an increase in literacy and numeracy. This suggests that changes in project designs and /or delivery processes are needed to ramp up the levels of literacy and numeracy achieved by the girls supported, and to improve the cost-effectiveness of the support provided.

4 Conclusions

What was the impact of the SCW after three years?

Girls' learning levels were extremely low at baseline. SCW projects reached and benefited a substantial number of marginalised girls, but did not manage to impact on learning at a sufficient magnitude across all contexts. At endline, girls in the communities targeted by the SCW are still falling substantially behind international benchmarks of reading fluency. Despite the absence of similar benchmarks for numeracy, the results show that projects found it even harder to improve numeracy than literacy skills.

Our analysis of EM data, which is intended to capture the effects of SCW projects as a whole on their target communities, does not demonstrate any significant impact over and above the control group in any of the main GEC outcomes (enrolment, attendance, literacy and numeracy). This implies that the **overall impact of the SCW on target communities has not been sufficiently large and consistent enough across** the different project contexts. Projects' results confirm this finding. SCW learning targets were typically defined as an improvement of 0.5 SD over the three-year implementation period of the GEC. Although these were not highly ambitious, eight out of 14 SCW projects met their literacy targets between baseline and endline, while only four met their numeracy targets. This does not mean that projects did not have a positive impact on their target populations. In terms of the number of marginalised girls supported by the GEC with improved learning outcomes[111], the FM reports[112] that the SCW 'substantially met' its learning outcome target supporting 577,618 girls, which represents 88 percent achievement against the endline target. But the magnitude of many projects' effects on girls' learning was not sufficient to achieve their targets. This is most likely due to a combination of issues relating to projects' design and delivery processes, including the time lag between some types of interventions and observable changes in girls' learning, and challenges in measuring girls' progress.

At endline, **girls' oral reading in targeted communities is still lagging considerably behind international benchmarks**. A reading fluency of 45 words-per-minute is considered a key threshold for basic comprehension in English. According to international benchmarks, children should achieve this at the end of the second year of their primary education. On average, girls across the SCW's target communities reached this threshold during their sixth year of primary education. This suggests that girls spend most of their time in primary school learning to read rather than reading to learn. Girls will not be able to progress effectively with their schooling as they get older and move up school grades if they are not reading sufficiently fluently to be able to understand what they are being taught.

Although improvements in literacy and numeracy scores seem largely correlated, **SCW projects found it harder to improve girls' numeracy**, as shown by the small number of projects that met their numeracy targets. This might be explained by the need for teachers having specific "hard" skills to teach numeracy that are not necessarily required to teach literacy[113]. Poor teaching, due to a lack of resources and under-qualified teachers, has been shown to be one of the most important barriers to girls' education across the SCW. Our analysis shows that higher attendance in school is correlated with improvements in literacy to a much larger extent than in numeracy, which suggests that the way numeracy is taught in schools may not allow girls to progress fast enough. This implies that **improving numeracy may require different types of interventions** than improving literacy, including potentially higher investment and a greater focus in training teachers how to teach numeracy.

---

[111] Defined by PwC for logframe reporting purposes *as improvements in literacy and numeracy by GEC-supported girls that exceed improvements by non-GEC supported girls, to an agreed target*.

[112] PwC (August 2017) "Girls' Education Challenge Programme Completion Report – Reporting Period 16 April 2012 – 30 April 2017"

[113] Atweh et al (2014) highlight that for early numeracy teaching, teachers must develop connected understandings of early numeracy content and problem solving, knowledge of the progression of numeracy learning development, syntactic knowledge related to problem solving as well as pedagogic content knowledge and skills related to early numeracy teaching. Additionally, fluency is considered essential for numeracy development, as a lack of fluency can impede conceptual understanding by taking up too much working memory and attention from thinking about the bigger mathematical picture. See *Bill Atweh, Arindam Bose, Mellony Graven, Jayasree Subramanian, Hamsa Venkat (2014). Teaching Numeracy in Pre-School and Early Grades in Low-Income Countries. GIZ.*

Being-in-school outcomes (enrolment and attendance) seem to have remained constant overall since baseline. For attendance, this may mostly be due to the challenges of accurately capturing relatively small changes. Enrolment rates were already high in some contexts, but in others a large portion of girls are still excluded from the schooling system.

Like literacy and numeracy, the EM's data does not show significant improvements in enrolment and attendance across the SCW's target communities, although some evidence suggests that projects seem to have been able to slow down drop-out rates of older girls. Attendance rates seem to have remained constant overall since baseline, with girls attending slightly more than four days out of five on average. However, the EM's and projects' research faced significant challenges in accurately measuring attendance, mostly because of the poor quality of school records and registers. Capturing changes in **attendance requires regular follow-ups and continuous measurements, and as such may be better captured as a monitoring indicator** rather than as part of an impact evaluation. Enrolment rates proved to be relatively high at baseline across the SCW, although our analysis showed that the **population of out-of-school girls is highly heterogeneous across contexts**. In some project areas, out-of-school girls are older girls and dropped out after having completed, or almost completing, a cycle of primary education. In other project areas, most out-of-school girls have never been enrolled. This suggests that improving the learning of these girls requires different type of interventions: on the one hand, to reduce drop-out rates (which can be achieved by improving attitudes and /or providing scholarships or bursaries to girls to help them transition to secondary school), or on the other hand, supporting or recreating girls' education environments, either at school or through extra-curricular activities.

Like out-of-school girls, some subgroups may face specific challenges and varying levels of marginalisation across contexts, but unfortunately none of them were tracked as specific cohorts, which makes it impossible to measure the impact of SCW on their outcomes.

The effect of the SCW on specific subgroups of girls within the projects' target populations is unclear. Although there is evidence from EM data that some of the activities were successfully targeted to girls who were the most in need, this does not seem to have translated into better learning gains for these girls. The absence of this effect is largely because **SCW projects usually considered their target population as a whole and did not make any distinction between the varying needs of different subgroups** within this population. When they did so, projects did not track the outcomes of these subgroups specifically. Even when subgroups can be identified in projects' data, associated sample sizes are usually not large enough to draw any meaningful conclusions on their learning progress. Besides, there is no consistent definition of subgroups across the SCW and projects often ended up using their own definitions, which makes any window-level comparisons more difficult.

What worked well and less well in the SCW?

Economic interventions have helped girls enrol and stay in school, but the assumption that this in itself will lead to improvements in girls' learning is highly dependent on the quality of teaching.

Economic interventions and in-kind support helped to reduce the cost of schooling and led to improved attendance among girls. Scholarships, bursaries and stipends all helped to offset the costs of education, and enabled households to overcome difficulties in affording school fees and sending their girls to school, while sanitary wear and menstrual supplies provided by projects have encouraged more regular attendance. Increasing income through loans, savings and cash transfers improved caregivers' willingness to spend more on girls' education, although there is little evidence that girls' learning outcomes improved as a result. Generally, projects provided little evidence of the links between changes in a particular barrier (at the output level) and the contribution this makes to improving education outcomes, in particular girls' literacy and numeracy. This is particularly the case for economic interventions that are designed to have an indirect effect on learning – they are dependent on the assumption that the school learning environment is good enough to improve girls' learning as a result of girls spending more time in school[114]. This is a big assumption for those projects that are not intervening to improve the learning environment,

---

[114] This has been debated over the past 20 years with mixed evidence. Advocates of shift schooling argue that lengthening the school day is an inefficient use of limited resources, as increases in instructional time have little effect on achievement (between 0.009 and 0.12 standard deviations) (Bray 2008). Nonetheless, studies reviewed by Orkin (2013) find that instructional time has positive effects on mathematics, with smaller or no effects on reading or language (Lavy 2012; Bellei 2009; Sims 2008; DeCicca 2007; Marcotte 2007; Brown and Saks 1987).

and one that needs to be measured and routinely monitored to ensure that their theory of change holds true, especially in the poorest contexts and where school education systems are particularly under-resourced and under-developed.

There is evidence that in certain contexts bursaries and stipends may be linked to improvements in learning outcomes as an indirect result of girls being able to enrol and attend school more than they otherwise would have[115]. Camfed (Zimbabwe and Tanzania) was able to demonstrate that the financial support they provided through bursaries led to higher attainment in English over the life of the project. In part, this may be explained by the education systems in these countries being sufficiently developed and resourced to support the assumption that retaining girls in school would in itself lead to improved learning.

Economic interventions have enabled poor households to send their girls to school, but there is a limit to the impact that SCW projects (mainly working at the local level) can have on households for whom poverty above all else is the binding constraint to girls' education.

The EM's and projects' evidence suggests that generally households have positive attitudes towards girls' education, but they are unable to send girls to school because they barely have enough income to meet more than their basic needs. Scholarships, bursaries and loans provide households with supplementary income that enables them to send girls to school. Projects have reported though that households' situations are changing as a result of drought, violence and economic decline that have reduced households' incomes. Some projects have also reported that the cost of schooling has increased. As girls age, there are strong expectations in many countries that girls will spend more time at home doing household chores. This is exacerbated where caregivers are working due to economic decline and need older girls to take on caring responsibilities in the absence of affordable childcare, or care for elderly or sick family members.

Projects have not been able to reduce the amount of time girls spend on household chores. Poverty is a structural barrier to girls' education. It underlies a range of different demand-side factors that prevent girls from accessing school and learning. For instance, our evidence suggests that projects have had limited success in reducing the effects of early marriage and pregnancies on girls' education. We found from our thematic research that early marriage and pregnancy are frequently used by families and girls themselves as a strategy for alleviating poverty. In some contexts, early marriage is the result of pragmatic decisions made by households who lack adequate livelihoods and as a result perceive that they have little choice. In the context of barriers driven by poverty, SCW projects have made a difference, but there is a limit to the impact they can have on households for whom poverty, above all else, is the binding constraint to girls' education.

Teaching quality and methodologies have improved, but systematic and sustained monitoring is needed to inform corrections to the delivery of projects' interventions and to provide a greater depth of understanding about how, why and to what extent girls are learning effectively as a result.

School-related factors, together with poverty-related factors, remain one of the biggest barriers to girls' education. All SCW projects addressed school-related factors within their theory of change as key barriers to girls' education, tackling: the quality of school facilities and sanitation; the availability and quality of teachers and teaching materials; classroom infrastructure; and school governance.

Teaching quality and methodologies have improved. Projects' endline reports found improvements in the use of participatory teaching practices and pedagogies since midline. Improved teaching methods helped engage girls in class, but frustratingly very few projects could provide any evidence that different types of improvements in the quality of teaching led to improvements among beneficiary girls. A lack of evidence does not necessarily mean a lack of impact. We expand on the broader methodological issue of projects not sufficiently evidencing the effectiveness of different types of interventions below. But in this instance, it may be that projects found it particularly difficult to measure the link between teaching and learning. Given its importance and potential impact on girls' learning in the classroom this seems to be a substantial gap in projects' monitoring and evaluation

---

[115] The relationship between bursaries and learning is supported by Glewwe and Muralidhan's (2015) and Unterhalter et al's (2014) review of five studies – conducted in Malawi, Benin, Kenya and China – that used RCTs to show that scholarship programmes led to improvements in learning outcomes across these different contexts. (Glewwe, P. and Muralidhan, K. (2015) Improving School Education Outcomes in Developing Countries: Evidence, Knowledge Gaps, and Policy Implications. RISE Working Paper.) (Unterhalter et al (2014) Interventions to enhance girls' education and gender equality. Education Rigorous Literature Review.)

frameworks. Similarly, it is unclear whether addressing girls' low self-esteem, aspirations and confidence leads to improvements in learning when these activities are not directly targeting improvements in literacy and numeracy. The theory of change is unclear in this respect and evidence of the links to education outcomes is lacking.

Projects need to systematically measure and track the quality of teaching and the effect on learning in the classroom to establish how effective projects have been and to help explain why, for example: some projects reported that teachers could not always recall the specific training they had received; and why despite the use of gender-responsive and participatory pedagogies, there was still evidence of negative treatment of girls in class among male teachers. Systematic and sustained monitoring of what is happening in the classroom is needed to inform corrections to the delivery of projects' interventions, and also to provide a greater depth of understanding about how, why and to what extent girls are learning effectively or not under different types of classroom conditions.

A range of school-based interventions, such as infrastructure improvements, learning materials, and literacy and numeracy clubs have improved learning, helping girls make substantial learning gains, but these gains have been constrained by significant systemic constraints in the supply of education.

Interventions that support girls directly with their learning have improved their ability to perform well in literacy and numeracy. Direct support with homework, reading and maths exercises has helped girls grasp difficult concepts and improve their confidence in their ability. Learning guides and books (i.e. materials) are enabling girls to study at their own pace. Interventions such as child-centred teaching methods, reading clubs and study guides are all helping girls receive the direct support they need to learn effectively. There have also been widespread improvements in classroom infrastructure, in particular sanitation facilities, which has encouraged girls to attend school more regularly.

However, projects have reported that the benefits realised from improvements in infrastructure and educational resources have been constrained by structural factors that are difficult to tackle. For example, several projects reported persistent concerns about the adequacy of new sanitation facilities that did not have access to water. The availability and supply of learning resources remains insufficient in several project areas and threatens to undermine the enrolment outcomes that have been achieved, particularly for the most marginalised beneficiaries. Projects have not been able to address the problem of schools being a long distance from where marginalised girls live. Problems with the supply of education infrastructure, learning materials and teachers are all symptomatic of education systems that lack the capacity and resources to provide a consistent quality of education that marginalised girls need. This is a systemic issue. SCW projects are generally tackling these issues at the school level where these symptomatic effects on marginalised girls are most profound, which is understandable. This is the point of intervention where SCW projects are able to provide immediate support to girls most in need and where they can achieve the required results with the time and resources available.

Projects are working at multiple levels of the education systems in which they are embedded. But there is little evidence that a holistic diagnosis of the education sector and the government and non-governmental systems that drive it is informing projects' intervention strategies beyond the relatively short three-year life of the projects. The SCW projects' evidence suggests that their achievements are being constrained by a severe lack of resources and capacity in the education systems in which they are working. As many SCW projects transition to the next phase of the GEC, it is timely for them to reflect on the evidence gathered so far and reassess the capacity of education systems to supply: teachers; learning materials; schools; classroom infrastructure; safe water; food; electricity; governance and management; and extension services for the most educationally marginalised. All of the SCW projects work with and in education systems that lack resources and are under-developed to differing degrees and in different ways. Projects need to use their data and evidence to reassess the effectiveness of each of their interventions in the context of the institutions in which they are embedded. If they do this, projects will then be in a better strategic position to pre-empt and avoid situations where, for example, new sanitation facilities lack water, or where projects are confronted by severe systemic problems that they are not able to address effectively.

Measuring the impact and effectiveness of the SCW

We need to know more about the characteristics of different subgroups of girls to fully understand what works, for whom and under what different types of conditions.

From an equity perspective, there is a lack of evidence about which groups of girls benefited the most from the SCW, how, why and with what effects on the problems they faced in accessing education. Throughout the evaluation process, from baseline to endline, it is clear that the factors that marginalise girls are context-specific within the same communities, schools and even classes in some situations. The current approach to longitudinally tracking a single cohort of girls in treatment and control groups treats them too much as homogeneous groups, which they are not. It is critical that we are able to understand and evidence the distribution of effects that projects have on different subgroups of girls. Projects' achievements should be evidenced by not only the impact on their overall target population, but also on these different subgroups. To do this, projects need to systematically and consistently collect more data about the specific characteristics of their target groups that goes beyond only collecting data on grade and age. Some projects currently do this more than others.

Tracking statistically significant subgroup samples for research purposes can be prohibitively expensive. However, for design, implementation and impact purposes, as much as for evaluation purposes, it is important that projects find cost-effective ways of identifying, diagnosing and addressing the different needs of the subgroups of marginalised girls that they are supporting. This can be achieved by making sure that household data is systematically and consistently linked to education outcome data across the SCW. What happens at home and in communities can have significant effects on what happens in the school, and the extent to which girls are able to learn and progress.

This is particularly important in the challenging and dynamic nature of the contexts in which projects work. For several projects, a range of external factors led to substantial changes in the problems their target girls faced – for example, the effects of: drought in Ethiopia, Zimbabwe and Somalia; the Ebola outbreak in Sierra Leone; and violence and political unrest in Afghanistan, Mozambique, Kenya and Ethiopia. At times, these factors also contributed to physical changes and attrition in projects' target populations and as a direct result of the characteristics of the subgroups they were supporting. For project design and implementation purposes, as well as monitoring and evaluation purposes, it is critical that projects have sufficient systems in place to track and identify these changes when they happen.

Finally, to conduct an accurate, meaningful and useful value for money (VfM) assessment and cost-effective analysis, data about the varying degree to which different groups of girls are marginalised is required. Particular factors that marginalise girls from education and make them harder to reach can have a significant effect on the cost of intervening to support them, for example, supporting mobile pastoralists in rural and remote parts of the Afar Region in Ethiopia. For our VfM assessment, by necessity we assume that the level of effort necessary to increase the learning of girls is the same across all projects' contexts. But to distinguish what type of interventions have been more or less effective in improving girls' learning outcomes (and at what cost), we need more information about their situation and characteristics.

The lack of evidence of the effectiveness of different types of interventions in project reports hindered our capacity to evaluate what worked well (or less well) in improving girls' learning outcomes.

While projects reported the effects of different interventions on barriers to girls' education, detailed and reliable evidence linking these changes to attendance or learning outcomes was often missing. Projects need to more systematically capture and present data in their reports that shows how, why and to what extent their activities contributed to reducing barriers to girls' education, and how these changes in barriers link to learning outcomes. Projects frequently failed to evidence the *intermediate steps* between outputs and outcomes. For example, SCW projects have reported an improvement in the quality of teaching. But it is not at all evident that girls' learning has improved as a result of this. Projects delivering interventions that were designed to have an indirect effect on girls' learning outcomes particularly struggled to evidence the theory of change between reducing particular barriers, like girls' self-esteem or empowerment and the effects of these changes on their learning outcomes.

Girls are typically marginalised from education as a result of multiple factors on both the supply and demand-side. These factors do not have the same effects on girls being able to access a quality education. Likewise, different types of interventions do not have the same effects on girls' situations. Holistic approaches involve different types of interventions that work in different ways, leading to different types of changes for girls. It is important that projects are able to understand and evidence in a more granular way what specifically worked well (or not) to inform course corrections, but also at the end of the project to identify which activities should be (or shouldn't be) sustained. This type of analysis was missing in projects' endline evaluation reports, and in particular in their sustainability assessments.

Greater oversight, more independent primary research and a more prescriptive approach is needed to support project evaluation to improve consistency and comparability in learning across the SCW.

The type and quality of quantitative data collected and submitted by each project varied greatly. Learning outcome test data was not consistently linked to the household survey and demographic data that was collected. Qualitative analysis presented in projects' endline evaluation reports was generally poor, providing limited explanations about how and why different types of interventions worked, for whom, in what contexts and with what effects. Despite a large investment of time and effort by the FM in particular, projects' evaluations generated limited evidence in terms of the causal relationships between barriers to education and girls' learning. This makes it very difficult for the EM to draw comparisons across the window and make generalisable observations about what has worked well or not. Without the benefit of independent qualitative research and analysis, it is difficult to establish the validity of the qualitative findings presented in projects' reports. This would have been strengthened by additional project-specific primary qualitative research undertaken by the EM to help validate the findings that were reported. If external evaluators are conducting project evaluations in GEC-T, then it seems clear that a more prescriptive approach is needed with greater technical oversight to deliver improvements in data quality, analysis, reporting, learning, consistency and comparability across the programme.

Sustainability and value for money of the SCW

Without the support of the successor to the GEC programme, GEC-T, it is highly unlikely that many SCW activities would have continued. There is little evidence that this will happen or could happen given the severe lack of resources available at all levels of the education system.

Without the support of the successor to the GEC programme, GEC-T, it is highly unlikely that many SCW activities would have continued as they stand now and very few would have continued beyond the immediate short term. This is perhaps unsurprising. The business case recognises that for all windows (not just the SCW) there was a risk that there would be a trade-off between sustainability and achieving short-term results. Many of the situations in which SCW projects are working are extremely challenging, where girls and their households face severe poverty and government education systems are woefully under-resourced. In these conditions, projects have had to work closely with the schools, communities and government officials and agencies in order to simply get their projects off the ground and to deliver results within a three-year implementation timeframe. In response, there is strong evidence of stakeholders' commitments to actively engage and support the activities projects delivered. This is evident in the match funding raised across the SCW exceeding its target by 160 percent to lever in £18.2 million against a target of £11.4 million.[116]

At the outset, projects were asked to report on the mechanisms that had put in place to enable marginalised girls to complete a full cycle of education. Projects' endline evaluation reports focused on the extent to which these mechanisms could continue to support girls without further support through the GEC. The FM reports[117] that all projects have put these mechanisms in place. At this stage, many projects' sustainability strategies are reliant on communities, schools or government ministries taking responsibility for continuing activities without much further support or funding. There is little evidence, however, that this will happen or could happen, given the severe lack of resources available at all levels of the education system and within communities facing the same economic pressures as parents.

---

[116] PwC (August 2017) "Girls' Education Challenge Programme Completion Report – Reporting Period 16 April 2012 – 30 April 2017"
[117] Ibid

In the short-term, the SCW has delivered improvements in girls' education, but the potential sustainability of these are severely constrained by the structural and systemic nature of barriers that projects are trying to address, and that have the greatest effects on girls' education.

At midline, we reported that projects planned, mobilised and implemented their sustainability strategies too late in their project implementation periods. There is no evidence to suggest that this conclusion does not hold true now at endline. Despite this, even if projects had planned and implemented their sustainability strategies earlier, it is clear now that some barriers to girls' education are extremely difficult to address without further financial support from outside local or national education systems.

Scholarships and bursaries, for example, are designed to enable poor families to send their girls to school. If parents are fully aware of the benefits of girls' education and have a positive attitude towards sending their girls to school, yet the only reason they do not is because they are so poor they cannot afford the direct and /or indirect costs of schooling, then someone else has to pay. Local school-based education projects in the SCW are typically not equipped to deliver sustainable solutions that address poverty and vulnerable livelihoods as a means of delivering statistically significant improvements in learning outcomes. Bursaries and scholarships solve a lack of household income in the short term but they are not sustainable, especially in the context of under-developed education and social welfare systems that are not able to cope with widely differing levels of incomes, sometimes within the same communities. This is why, at midline, a key conclusion was that projects had to join up more with other programmes and initiatives that were better equipped to address big structural problems such as poverty, livelihoods, water supply, health and sanitation, food security, etc. At endline, the evidence suggests that this is still an important finding, particularly with regards to the supply of safe water to schools, which for some projects has severely constrained the benefits from constructing new toilet facilities and girls' capacity to concentrate and learn while in school.

Similarly, the long-term benefits of training teachers and the evident improvements in teaching have been compromised by high levels of teacher transfers and a lack of resources in schools and education authorities to be able to maintain the training that teachers need to continue their professional development. These are structural issues that to a large extent are outside the control of the type of projects delivered through the SCW. Under-developed education systems lack the resources and capacity to quickly adapt the ways in which they work and resource education within a relatively short timeframe (three years). The underlying causes affecting the supply of teaching and school infrastructure are systemic, and as such take much longer and much larger investments to address in a sustainable way than SCW projects could reasonably afford.

The SCW has improved the literacy and numeracy outcomes of a large number of marginalised girls living in challenging conditions while simultaneously facing under-resourced school and education systems. Projects, though, need to significantly adapt their designs and delivery processes to increase their effects on girls' literacy and numeracy, to enable them to effectively progress through their education.

The SCW has improved the literacy and numeracy outcomes of a substantial number of marginalised girls living in challenging conditions while simultaneously facing under-resourced school and education systems. At the start, girls' learning levels were appallingly low. They are still far too low for girls to progress through their education at the right levels given their age and grades. Despite the achievements of the SCW, at the end of the GEC, projects' interventions have not delivered the magnitude of improvement in literacy and numeracy levels, despite demonstrating some effects that have benefited a reasonably large number of marginalised girls. Improvements in the intensity and magnitude of the learning gains achieved by GEC projects would improve the cost-effectiveness of the most effective interventions. It may be that more time is needed for some types of interventions to deliver their full potential in terms of increases in literacy and numeracy. However, the findings from the evaluation overall suggest that significant changes in some project designs and /or delivery processes are needed to increase the levels of literacy and numeracy gains in order to improve the effectiveness, impact and overall value for money.

5 Recommendations

The following recommendations have been developed with a view to specifically informing the next phase of the GEC, which was launched in June 2017, and consists of the GEC 'Transitions' Window and the 'Leave No Girl Behind' Window. These recommendations are aimed at DFID. However, the actions required to deliver them potentially involve DFID, the FM, grantees and the future independent Evaluation Service Provider.

1. **A more intensive focus and effort is needed to improve the magnitude of projects' effects on girls' learning levels.** The SCW has reached and benefited a substantial number of marginalised girls. However, girls' learning levels are still too low for their respective grades, hindering their effective progress through their education. SCW projects need to increase the magnitude of their effect on the learning levels of the cohort of girls supported in the next phase of the GEC. This recommendation builds on that made at midline, which highlighted the need for **projects to focus on interventions that have proven to (or are most likely to) have the greatest impact on girls' learning levels within the required timescales.**

2. **Projects need to develop realistic sustainability strategies that are based on a diagnosis of the type of barriers to girls' learning that can be sustainably resolved within project timescales, compared to large (often underlying) structural or systemic causes (e.g. livelihoods /poverty, school infrastructure, supply of teachers) that require much larger and longer-term investments at a systems level.** Many projects' sustainability strategies were overly reliant on communities, schools or government ministries to continue activities once projects had come to an end. In part, this was because sustainability strategies were developed too late in projects' lifecycles, with the underlying barrier to sustainability being a pervasive lack of resources within the target communities, schools and education systems. Sustainability strategies should explicitly recognise the trade-offs between addressing symptomatic effects to achieve short-term learning gains, versus what might be considered the longer-term objective of addressing structural causes of poor education outcomes for marginalised girls to achieve sustainable changes in learning.

3. **Projects need to develop interventions specifically designed to improve girls' numeracy.** Overall, the SCW has had less of an impact on numeracy than literacy. The endline findings suggest that girls improve less in numeracy than in literacy when they are in school, and that in some contexts, teachers find it more challenging to teach numeracy than literacy. This implies that improving numeracy requires a more intensive effort and different types of intervention compared to literacy. **Projects should design interventions that are based on a clear diagnosis of why targeted girls struggle to improve their numeracy compared to literacy, including analysis of the challenges schools and teachers face in being able to effectively teach numeracy.**

4. **Attendance should be monitored as an intermediate indicator of improvements in learning.** Attendance proved very difficult to track and measure across the SCW. While household-based measures are not precise enough to capture small changes in attendance levels, the validity of school-based measures is often limited by the low reliability of school registers and records. Evaluation surveys may therefore not be the most appropriate tool to capture attendance data. Measuring attendance requires adapted strategies and tools, as well as the involvement of staff in the field to conduct regular follow-ups and spot checks throughout the course of the programme. **Attendance indicators should therefore form part of projects' monitoring strategies as well as being included as intermediate outcomes that support their evaluation strategies**.

5. **Projects need to demonstrate a better understanding of how to support different needs of out-of-school girls.** Out-of-school girls were often treated as a homogeneous group, under the assumption that all girls who are out-of-school have similar educational needs and require similar types of intervention. However, the endline evaluation findings suggest that out-of-school girls are a heterogeneous population - formed, on the one hand, of girls who have mostly been excluded from the schooling system (either because they dropped out early or because they never attended school), and on the other hand, of older girls who dropped out after completing a cycle of primary education. **Projects need to explicitly recognise the varying needs of out-of-school girls**

rather than assuming they are a homogeneous subgroup. The former need to access learning environments, sometimes for the first time, where they can learn effectively, whereas the latter need to be re-engaged by schools and education systems and supported to overcome barriers that prevent them from staying in school and transitioning through education and schooling phases.

6. **Projects need to better evidence the distribution of their effects on different groups of girls.** More systematic subgroup analysis is required to understand the effects that projects' activities have on different groups of girls affected by different types of barriers. **Projects need to be able to identify and track the extent to which specific subgroups are marginalised from achieving education outcomes. This includes gender comparisons of girls and boys.** At the design stage, projects should collect evidence of which socio-economic subgroups are the most disadvantaged in their operating context, and in particular, whether girls are significantly disadvantaged compared to boys. Once subgroups have been defined, their outcomes should be tracked throughout the course of the programme (which includes collecting samples of sufficient sizes to be representative). For this, household data (which usually includes personal, social and economic subgroup characteristics) need to be systematically linked with learning outcome data. Projects' achievements should then be measured as the impact achieved by projects not only on their overall target population but on these different subgroups.

7. **Projects need to better evidence how changes in a particular barrier will translate into better learning outcomes.** This is particularly important for barriers that are not directly related to learning, such as economic and poverty barriers. The link between barriers and learning outcomes and the intermediate steps between them need to be explicitly measured to ensure that theory of change continues to hold true. Projects also need to monitor changes in external conditions, which can cause changes in barriers. This is particularly true for economic conditions whose deterioration can impact across a range of factors both on the supply and demand-side, eventually threatening the validity of the whole theory of change.

8. **Projects need to systematically monitor changes in the quality of teaching to ensure that activities aimed at improving teaching quality are on the right track**, and to inform corrections if this is not the case. Projects have largely been unable to measure changes in the quality of teaching and generally assumed that activities to improve teaching and pedagogy (such as teacher training) had a positive effect on learning, without testing this assumption. The evidence suggests that this assumption did not always hold true. More generally, it is crucial that projects are able to track and monitor the effectiveness of different types of activities, not only to inform course correction during implementation, but also to evidence what activities worked well or not, with a view to identifying which ones should be sustained (or not) after the end of the programme. Again, this is particularly important for activities where the theory of change and causal link to improved learning is unclear, as with activities designed to improve girls' confidence and self-esteem.

9. **It is crucial that projects conduct context analysis as an integral part of their design and M&E process. Contextual factors have had a significant effect on projects' success, and projects need to take into account the effects of wider structural constraints that can threaten the validity of their theory of change and undermine their achievements**. This includes carrying out a political economy diagnosis of the education system to ensure that the theory of change does not overestimate the system's capacity to supply basic educational needs (such as teachers; learning materials; schools; classroom infrastructure; safe water; food; electricity). Projects will then be in a better position to anticipate systemic constraints that can hinder the effectiveness of their interventions, such as the lack of running water when building toilet facilities.

10. **Projects need to improve their capacity to demonstrate and report** *how* **and** *why* **their interventions have worked well or less well, with what effects on learning, for whom and under what different type of conditions.** As recommended previously, a greater focus on identifying and tracking the characteristics of different subgroups will help projects do this. Similarly, identifying and evidencing the intermediate steps between outputs and outcomes as an integral part of their M&E frameworks will also help explain how and to what extent some interventions are having a greater effect on learning than others. Projects should also

systematically report on the scale of activities delivered, and the process through which they were delivered, to enable better evaluation of the effectiveness of their interventions. The Terms of Reference for projects' external evaluators should be more prescriptive, including requirements to: use a standard household survey template; link household survey data with learning outcome data; and collect specific disaggregated data about key subgroups. Furthermore, the FM could consider requiring projects to provide qualitative data, for example in the form of redacted transcripts (or notes) of interviews and focus group discussions. The intention is to improve the external validity of the quantitative and qualitative analysis and findings produced by external evaluators, and enhance the quality of learning from significant investments in these evaluations.