coffey
A TETRA TECH COMPANY

Midline Evaluation Report

Step Change Window

Final Version (March 2017)

Evaluation Manager Girls' Education Challenge (GEC) Fund

ORB International

RTI INTERNATIONAL

Midline Evaluation Report

Step Change Window

Final Version (March 2017)

Evaluation Manager Girls' Education Challenge (GEC) Fund

UK Department for International Development

Evaluation Manager Girls' Education Challenge (GEC) Fund

PO 5685

Partners

- RTI International
- Opinion Research Business (ORB)

Lead Authors

- Simon Griffiths
- Karolin Krause
- Florian Poli
- Maha Batran

Contributions by

- Evelyn Hytopoulos
- Paul Diegert
- Ella Page

March 2017

Coffey International Development Ltd

The Malthouse 1 Northfield Road Reading Berkshire RG1 8AH United Kingdom
T (+44) (0) 1189 566 066 F (+44) (0) 1189 576 066 www.coffey.com
Registered Office: 1 Northfield Road Reading Berkshire RG1 8AH United Kingdom
Registered in England No. 3799145 Vat Number: GB 724 5309 45

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

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	Female Genital Mutilation
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 Deviations
SEM	Structural Equation Modelling
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

In 2012, the Department for International Development (DFID) launched the £355 million Girls' Education Challenge Fund (GEC). The GEC sets an ambitious target of reaching one million marginalised girls by the time its first phase ends in April 2017[1]. The Business Case[3] for the GEC recognised that there was a lack of robust evidence about the benefits of focusing on girls' education. At the time, DFID, other donors and policy-makers did not know enough about how and why girls were marginalised in terms of their education and how best to intervene to bring about significant changes in the lives of marginalised girls. This was the underlying rationale for a fund that challenged the market to identify the causes of educational marginalisation, propose effective strategies for improving education outcomes for girls and rigorously collect evidence to prove and explain what worked well, why, for whom and under what types of circumstances.

DFID appointed Coffey, in partnership with RTI International and ORB, as the Evaluation Manager (EM) of the GEC. The EM is responsible for independently evaluating the overall effectiveness and impact of each funding window. The EM closely collaborates with the GEC Fund Manager (FM) to support projects in collecting data in line with the programme's evaluation requirements, and in reporting results with a maximum level of consistency across the fund.

The Step Change Window

The Step Change Window (SCW) is one of three funding windows of the GEC. The other windows are 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 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 midline evaluation. They are delivering holistic approaches that aim to address multiple barriers to girls' education at the level of the individual girl, the household, the community, and the school.

Purpose of the midline evaluation report

This midline evaluation report follows the SCW Baseline Report that the EM published in January 2015. The EM plans to complete an endline evaluation of the SCW later in 2017. This evaluation covers the first two years of the projects' three-year implementation period in the first phase of the GEC. The purpose of this midline evaluation report is to provide reliable evidence of the programme's impact on being-in-school and learning outcomes, effectiveness and sustainability that DFID, the FM and projects can use to inform further development of activities funded by the GEC – in particular, the next phase of the GEC (an additional four years are planned after the current phase finishes in 2017[4]). This report should also generate transferable learning for a wider audience including donor agencies, governments of GEC countries, and other policy-makers.

Evaluation and research approach

The EM produced the GEC Evaluation Strategy towards the end of 2012. The strategy set out a two-pronged approach to evaluating the SCW. The EM focuses on the effectiveness and impact of activities delivered by all projects across the window on the catchment communities they targeted. For that purpose, we selected a random sample of girls who live in the catchment areas of the treatment and control schools for all fourteen SCW projects and were aged 5-15 at baseline and are now 7-17 at midline. Our samples are designed to be representative of the general SCW target population as a whole, which means they are 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

[1] Girls' Education Challenge: http://devtracker.dfid.gov.uk/projects/GB-1-202372
[3] GEC Business Case v.3, September 2011
[4] GEC Business Case v.3, September 2011

evaluators to design their own separate research samples that would be sufficiently representative and large enough to enable individual projects to measure their 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. The difference in perspectives between the impact evaluation approach taken by the EM and those taken by projects is important. 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 midline evaluation consisted of: 6,084 longitudinal household surveys and literacy (EGRA) and numeracy (EGMA) tests; 2,853 school visit surveys; and 744 in-depth qualitative interviews of caregivers, teachers, and community leaders. In addition, a school-based assessment of learning outcomes and conditions in intervention and control schools was conducted in five project areas in DRC, Ethiopia, and Kenya. It involved 406 classroom observations, 427 teacher surveys and 7,237 learning tests with boys and girls – this data was specifically used to evaluate gender disparities in literacy and numeracy over the course of the project. The EM also conducted reviews and meta-analysis of project midline evaluation reports, datasets and outcome spreadsheets submitted by SCW projects to triangulate the analysis of the EM's quantitative and qualitative data.

Key findings

Reach and equity

SCW projects have reached girls who are highly disadvantaged, showing high social, economic and environmental markers of marginalisation. This finding is also supported by findings from the EM's process review of the GEC[5]. We found at the midline stage that girls across the SCW target communities differ substantially from one context to another in the ways that they are marginalised and the extent to which they are marginalised from education. At the start of the GEC, DFID decided not to prescribe the type of factors or underlying causes that marginalised girls from education. This was appropriate because there was a lack of evidence about what caused girls' marginalisation and it was clear that the factors driving complex problems would vary from one to context to another. For this reason, applicants were required to explain how and why the girls they were targeting were marginalised from education[6]. Most projects used geographic criteria to target marginalised girls, while others used ethnographic targeting criteria. Generally, a lack of specific project definitions of marginalisation has made it difficult to assess the extent to which projects effectively impacted on the barriers that marginalised their target girls compared to other girls and boys in their communities.

Impact on learning

DFID's Annual Review of the GEC[7] reports that "the SCW exceeded its logframe target for the number of girls with improved learning outcomes of 449,300 by 23% (552,793) – this includes both girls who met the set learning targets and those who outperformed their peers but did not meet learning targets. Within this figure, 66% of the girls who improved their learning outcomes came from five projects – IRC (DRC); CfBT (Kenya); Camfed (Tanzania); BRAC (Afghanistan); and AKF (Afghanistan)". These project learning figures combine literacy and numeracy results.

For literacy, six out of twelve[8] SCW projects met or exceeded their midline literacy outcome targets according to their own impact evaluation data and reports. This demonstrates high levels of performance across these projects. Looking at the impact on wider SCW communities, the EM data shows no evidence of positive effects of SCW projects on the average reading fluency of girls in treatment communities compared to control communities. However, the SCW has had a positive statistically significant effect of four words per minute on the *median* reading fluency score. At baseline, across the EM sample more than half the children who were tested could not read a single word. Now, at midline after two years, half the girls can read at least 23 words per minute in the control group and 27 words per minute in the treatment group. Further analysis shows that the effect of the SCW on reading fluency has tended to focus around the middle of the score distribution (mid-performing girls who score around or slightly below the median) while the effect on numeracy has been the largest on lowest-performing girls.

[5] Coffey (2015), GEC Process Review Report
[6] Ibid
[7] DFID (2017) GEC Annual Review Report
[8] Plan (Sierra Leone) and Acted (Afghanistan) did not submit midline evaluation reports, although Acted recently submitted an endline report because their project has now come to an end.

The EM data also shows that the SCW had a net positive effect on literacy among out-of-school girls – their reading fluency improved significantly since baseline by 12 words per minute more than the control group.

Our analysis of gender disparity in learning from testing boys and girls in the DRC, Ethiopia and Kenya project areas confirms our baseline findings that in-school girls do not have systematically different levels of learning in comparison with in-school boys of the same grades. The gender gaps related to attendance and learning were relatively small at baseline for children in primary school. Girls in treatment and control areas showed greater improvements in literacy and numeracy scores than boys of the same age and school year. In other words, girls in this sample of primary school students are now overtaking boys across the SCW.

Projects' own midline evaluations reported that three out of eleven[9] projects met or exceeded their numeracy outcome targets at midline. Three other projects partially achieved (i.e. greater than 60%) their numeracy targets. Projects have not been able to improve numeracy to the same extent as they were able to for literacy, which according to DFID's Annual Review[10] "may be due to a lack of focus on numeracy teaching in class and poor confidence in teaching it". Using the EM data, we show no evidence of an overall increase of the average numeracy of girls in treatment communities compared to control communities. However, similar to literacy, numeracy improved significantly since baseline for out-of-school girls by 13 percentage points more than the control group.

Impact on attendance

The EM data shows that in-school attendance (and enrolment) rates have not significantly improved overall across SCW target communities compared to control areas. However, the difficulty of capturing small changes through household-based measures has meant we do not have conclusive findings at midline about improvements in girls' attendance at school. The FM also reported inconsistencies and overall poor quality of the measures of attendance used by projects. DFID's Annual Review of the GEC recognises that "attendance data should be interpreted with caution given inconsistencies in measuring attendance at the school level, unreliable school records which tend to inflate attendance levels, and restrictions/resistance to spot checks, especially in government schools." The EM acknowledges this and faced similar challenges during the EM in-school research.

Effectiveness of SCW projects' activities

Across the SCW and on the basis of the EM data, the majority of the most important barriers to girls' education identified at baseline in target communities have not changed significantly as a result of SCW activities. A few barriers, mostly related to community attitudes and caregivers' satisfaction with school facilities, teaching and school affordability have improved significantly across the SCW projects. It proved difficult to accurately capture changes in the quality of teaching and school infrastructure through quantitative surveys.

The activities which seem to have had the biggest impact are those which aimed at improving learning directly, such as special tutoring, help with school work and teacher training. Projects who invested the most in economic interventions are the same projects that managed to improve school affordability substantially. The provision of bursaries and in-kind support, as well as loans seems to have been an effective way of alleviating the burden of school-related expenses in some communities, but this does not seem to have translated into better learning outcomes yet.

Four projects experienced and reported a backlash and resentment from boys that might partly be explained by boys being unwilling to give up gendered privileges, but it may also reflect that boys have many similar needs to girls that are not addressed by the GEC. Another five projects reported backlash from community members because of activities targeting girls exclusively, and mentioned the negative impact this may have had on the education of both boys and girls.

Sustainability

The SCW exceeded its midline match funding target of £8.1m by generating match funding of £12.5m (a 54 per cent over-achievement). Projects report a limited amount of information about how match funding is used to the FM. So while the SCW has performed very well against its midline target, it is unclear how much of this match funding indicates initial stakeholder buy-in to project approaches at the start and /or a commitment by stakeholders to sustain project activities.

[9] Plan (Sierra Leone) and Acted (Afghanistan) did not submit midline evaluation reports, although Acted recently submitted an endline report because their project has now come to an end. CfBT (Kenya) numeracy data was not deemed valid for baseline to midline comparison.
[10] DFID (2017) GEC Annual Review Report

Several projects have demonstrated that they have effectively engaged and liaised with government in a variety of ways. However, many projects reported the difficulties that they experienced in collaborating with government and trying to get enough buy-in and ownership to sustain key activities beyond the life of the project. In the absence of government capacity and funding, many of the SCW projects' sustainability strategies appear dependent on continued community support. Although it is generally unclear how and why community-based support will succeed in sustaining the level of involvement and contributions needed to continue to deliver project activities and achieve similar changes and results. Generally, there is little evidence at the midline stage to suggest that the SCW projects will be able to sustain project activities by the end of the first phase of the programme in July 2017 without further external funding.

Key conclusions

The SCW has exceeded its logframe midline learning target. Six out of twelve projects met or exceeded their literacy targets while three projects met or exceeded their numeracy targets – some of these projects have achieved large learning impacts for their target girls. However, based on EM data, there is little evidence of an overall impact on girls' learning in communities targeted by projects. But the SCW has had a positive impact on out-of-school girls and those at the bottom end of the learning spectrum.

The EM data shows no evidence of an overall effect of SCW projects on the average reading fluency or numeracy of girls in treatment communities compared to control communities. This implies that, although some projects have met their learning target, the average impact achieved by SCW projects as a whole on girls in their target communities is not large enough yet to be captured by the EM sample. However, the EM data demonstrates that specific subgroups have significantly improved their literacy and numeracy as a result of SCW activities: mid-performing girls have improved their reading fluency more than in the control groups, as demonstrated by the 4 words per minute net increase in the median score. Girls with the lowest scores in numeracy have also increased their scores significantly. Besides, the EM data suggests that the SCW had a net positive effect on literacy and numeracy since baseline among out-of-school girls compared to the control group. Many projects established girls' clubs, tutoring, or alternative education facilities for out-of-school girls to help them prepare for re-integration into mainstream schools or to teach them basic numeracy, literacy, life and vocational skills. The positive results observed for out-of-school girls suggests that creating additional or alternative learning environments have been effective in the short-term. However, it may prove difficult for projects to sustain solutions that are developed outside established school systems, and to scale up these types of approaches with schools and communities who lack funding and resources.

Across the SCW and on the basis of the EM data, the majority of the most important barriers to girls' education identified at baseline in target communities do not seem to have changed significantly as a result of SCW activities. While there are potentially several factors constraining projects' ability to address barriers that are particularly difficult to overcome, it's not clear how and why projects have identified which barriers they should focus on to directly deliver changes in learning within relatively short timescales.

At the start of the GEC, SCW projects targeted a variety of different populations with different needs and starting points. Projects made trade-offs between the number of girls they aimed to reach, the degree to which girls are marginalised and the extent to which they would be able to deliver sustainable outcomes.

Our midline analysis of marginalisation shows that SCW target girls are often highly disadvantaged, but the type of factors that marginalise girls and their effects on education vary significantly from one context to another. Projects were required to analyse and articulate how marginalised girls differed from others in their target communities, what their specific education needs were, and what distance they must travel to significantly improve their learning outcomes. For those projects working with particularly marginalised groups such as disabled girls, young mothers, or pastoralist communities, achieving their learning targets may be more difficult than for those working with a population of in-school girls who are not necessarily more marginalised than the rest of their communities. It is clear from the EM data that projects' decisions at the start about who they target has involved making trade-offs between: the number of girls they need to reach and benefit; the levels of severity and complexity of problems that marginalise girls from education; and the extent to which underlying causes can be addressed to deliver sustainable outcomes in the longer-term.

SCW project areas are a crowded space. There is a lack of evidence that projects have mapped their strategic landscapes to identify potential partners and the best entry points. As a result, there appears to be little evident coordination or complementarity between GEC projects and non-GEC actors.

SCW communities are a crowded space for education programming and our evidence suggests that non-GEC activities are taking place in most treatment and control areas that address similar educational barriers. This may have led to "contamination" of projects' control groups and as a result their impact evaluation findings. At this stage of implementation, there is little evidence that projects have coordinated with other GEC projects and non-GEC actors for the purpose of joint programme planning, design or implementation. Only a few project midline evaluation reports explicitly mentioned the activities of other actors in GEC treatment areas even though the EM's qualitative data suggests that these were prevalent in most project areas. A key learning point from midline is that projects would have benefitted from conducting more comprehensive strategic landscaping before the start of implementation. However, the lack of a formal Inception Phase for the programme as a whole and the SCW[12] meant that there was little time and no explicit requirement for projects to map the strategic landscape and scope potential partnerships with government and other actors at the pre-baseline stage.

SCW projects provided little analysis of how different types of interventions have addressed barriers that specifically impact on girls' education more than boys. As a result, the extent to which projects have improved gender disparity in learning and attendance outcomes at midline is not clear.

Our analysis of gender disparity in learning from testing boys and girls in the DRC, Ethiopia and Kenya project areas confirms our baseline findings that in-school girls do not have systematically different levels of learning to in-school boys of the same grades. Furthermore, several projects experienced and reported a backlash and resentment from boys that may partly be explained by boys being unwilling to give up gendered privileges, but it might also reflect that boys have many similar needs to girls. These findings point to the importance of exploring gender gaps in context for the purpose of distinguishing between universal barriers to learning from gendered problems that affect girls more than boys. It is also critical to understand what effects gender dynamics have on girls' education immediately and how these evolve and take effect as girls grow older. This type of gender analysis was lacking in projects' midline evaluation reports. Testing boys in addition to girls was not a requirement for projects. As a result projects are not in a position to analyse where gender gaps exist and how they might have changed.

Sustainability needs to be locked into projects from the start and become an integral part of the problem diagnosis, project design, monitoring and evaluation processes to deliver best value for money.

Sustainability was prioritised at the start as a key logframe outcome that projects were required to deliver. However, in practice many projects started developing and implementing their sustainability strategies later on in the implementation process. At midline, projects now face a race against time to get enough stakeholder buy-in and develop enough community resilience and capacity to continue delivering the right activities in the right ways to continue benefiting the projects' target groups. Several projects engaged with government in a variety of ways, but many struggled to generate enough buy-in and ownership to sustain key activities beyond the life of the project. In the absence of government capacity and funding, many of the SCW projects' sustainability strategies appear dependent on continued community support. It is generally unclear though how and why community-based support will succeed in sustaining the level of involvement and contributions needed to continue to deliver the activities required to deliver similar results by the end of the projects in March /April 2017.

The SCW has demonstrated that rigorous experimental and quasi-experimental impact evaluations are possible even though they require expert resources and stable conditions that are not always available on the ground. Some projects' external evaluators struggled to produce good quality qualitative analysis and evidence of the strengths and weaknesses of specific project interventions.

The GEC's Evaluation Strategy requires a particularly rigorous approach to evaluation and data collection. The GEC process review[13] reported that the GEC had accomplished its objective of embedding robust M&E in a complex, large-scale, multi-country, multi-window programme. However, the impact evaluation process has met substantial challenges as a result of contexts marked by insecurity and instability, natural disasters, and migration, further challenged by situations where both projects and external evaluators have relatively little experience of implementing impact evaluation designs involving longitudinal surveys across treatment and control groups.

[12] Coffey (2015), GEC Process Review Report
[13] Coffey (2015), GEC Process Review Report

Future evaluation designs should take into account that it is very difficult to draw general conclusions about the effectiveness of multi-country programmes in which each project targets a distinct population and context with a distinctive theory of change. To produce sufficiently meaningful learning about what works in addressing key barriers in a specific context requires a clear understanding of intervention mechanisms and their specific theories of change at the project level. To do this, good quality qualitative evidence about what works, why, for whom, where and with what specific effects on learning outcomes is needed at the project level. Project-level evaluations would have benefitted from a more systematic assessment of the theory of change, by testing assumptions and hypothesis systematically, tracking whether intermediary outcomes have been achieved, and monitoring changes along the various steps in their results chains and impact logic.

1 Introduction

1.1 Background and context of the GEC

1.1.1 Context and rationale

Changes in the 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.[14] 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.[15]

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[16] and 3[17], 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[18].

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 263 million children and young people were estimated as being out of school by the end of 2014 – this includes 61 million children of primary school age (6-11 years), 60 million adolescent children of lower secondary school age (12-14 years) and 142 million children of upper secondary school age (15-17 years)[19]. In secondary school, the differences between boys' and girls' participation rates are significant. Large disparities exist within countries, with poor rural girls experiencing worse educational outcomes, even at the primary school level.

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[20] - approximately 250 million children cannot read, write or do basic maths, including more than 130 million children who do not have basic 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 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.[21]

It is estimated[22] 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.

[14] 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.
[15] 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
[16] 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.
[17] 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.
[18] EFA (2015), 'Global Monitoring Report: Education for All 2000-2015, Achievements and Challenges', UNESCO
[19] UIS and GEM (2016), 'Leaving no one behind: How far on the way to universal primary and secondary education? Policy paper 27 /Fact Sheet 37. Paris: Global Education Monitoring Report'.
[20] UNESCO (2014), 'Teaching and Learning: achieving quality for all'. EFA Global Monitoring Report, Paris: UNESCO
[21] 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.
[22] EFA (2015), 'Global Monitoring Report: Education for All 2000-2015, Achievements and Challenges', UNESCO

Persistent under-investment in education

A key rationale[23] 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 suffering the greatest decline. While total ODA rose by 11 per cent in 2013, aid to basic education declined by 7 per cent[24]. 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[25] 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 (FGM); 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.[26]

Extending DFID's Girls' Education Challenge Fund

The current GEC programme ends in July 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[27]. 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[28]. 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 who are 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[29] 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

[23] DFID (2012), Girls' Education Challenge, Business Case Version 4; London: DFID
[24] 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
[25] Coffey (2015) 'Baseline Report – Step Change Window' https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/425360/Step-change-window-full2.pdf
[26] https://www.unicef.org/emergencies/files/FINAL_PUBLIC_Consolidated_Netherlands_Donor_Report_2009.pdf
[27] 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.
[28] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/559123/leave-no-girl-behind.pdf
[29] Coffey (March 2013), Annex A in the 'GEC Evaluation Strategy'.

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, attendance, retention and learning in school and contribute to an overall impact of improved life chances for marginalised girls.

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[30] projects that were operating at the time of the midline evaluation developed holistic theories of change that aim to address multiple barriers to girls' education at the level of the individual girl, the household, the community, and the school. Some projects also work to improve the governance of girls' education issues inside and outside school.

Projects tackle these barriers through a wide range of interventions. The FM has produced an intervention mapping that is summarised in Table 1. 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.**

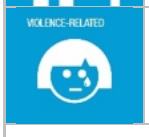
In addition, **some projects** carry out additional **activities to empower girls** and raise their self-esteem, to **tackle marginalisation** (e.g. by addressing barriers related to disability, or issues of cultural or linguistic exclusion), and to **reduce violence** in school or in the community.

A full overview of the activities that each project is delivering, and short project profiles are 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 midline evaluation is provided in Section 3.1.

Table 1: Overview of SCW project interventions

	BRAC Afg	AKF Afg	Acted Afg	IRC DRC	STC Eth	ChHp Eth	WUSC Ken	CfBT Ken	STC Moz	Plan Sie	RI Som	CARE Som	WV Zim	Camf Z-T
ECONOMIC	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
INFRASTRUCTURE & RESOURCES	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

[30] Please note that only fourteen projects were active at the time of the midline evaluation. The GEC grant held by BRAC for its project in Sierra Leone ended in 2015.

	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	✓	✓	✓	✓	✓	✓	✓	✓	✓	✦	✓	✓	✓	✓
	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✦	✓	✓
	✓	✓	✓	✓	✓	✓	✓	✓	✓	✦	✓	✓	✓	✓
	✓	✓	✦			✓			✓		✦	✦	✓	✓
	✦				✓	✓	✓	✦		✓		✦	✓	✓
	✦	✓	✦		✓	✓			✓		✓	✓	✓	✦

Note: ✓ 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.
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[31] 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.

Box 1: A note on how we refer to SCW projects throughout this report

For simplicity, readability, and consistency with the SCW baseline report 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 equally be recognised for their contribution.

The "Steps Towards Afghan Girls Educational Success" (STAGES) programme led by the Aga Khan Foundation (AKF) in Afghanistan for instance involves a range of partners such as Care International, Catholic Relief Services, Save the Children, Afghan Education Production Organization and Roshan Telecom who work hand in hand. CARE is building schools in the areas where there is a cluster of Community Based Education (CBE) classes that will become a government school after project handover and Save the Children and AKF are renovating and equipping existing government schools to ensure that they provide a stimulating learning environment.

More information on national and international project partners can be found in the projects' individual midline evaluation reports.

[31] 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.

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 level. This approach reflects the perceived wisdom at the start of the programme that to tackle barriers to girls' participation in and benefit from education a holistic approach was needed. Differences in approaches or intervention design 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 specific 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 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.

1.1.4 The GEC in the context of other girls' education programmes

The GEC was part of a global effort to achieve the MDGs 2 and 3 and is now contributing to SDGs 4 and 5. In light of declining ODA to education, DFID set out to focus its efforts on girls' education through the GEC with the ambition that it would have a catalytic effect on other international partners. While multilateral agencies, national donor organisations, and NGOs are implementing girls' education projects around the world, the GEC has been unique in its breadth of scope, level of funding, and programme design. Annex C provides an overview of other, similar girls' education programmes and how they differ from the GEC.

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 midline 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 average poor quality of project data and the insufficient level of consistency across them (inconsistent surveys, inconsistent sampling designs, inconsistent learning tests, etc.) this has not been possible either at baseline or at midline (see Section 2.5.2 about the limitations of SCW projects' data).

1.2.2 Purpose of the GEC midline 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. The GEC Knowledge Management Working Group led by the FM has a key 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 are a member of

this Working Group. In addition, the EM is currently developing, with DFID, a specific communication plan for the findings of this midline evaluation that will target key education partners of DFID, including UNICEF, UNGEI, USAID, GPE, UNESCO and the World Bank. Activities in the communication plan will be delivered in early/mid 2017 following the publication of this report.

This midline evaluation also serves an important accountability purpose by providing reliable information about the effectiveness and impact of the SCW projects two years into the three-year implementation period. It follows the baseline research that was conducted in 2013/14, and precedes the programme endline evaluation that will be completed by December 2017.

1.2.3 Scope of the GEC SCW midline evaluation

The midline 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, retention 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 midline evaluation focuses on changes in outcomes (i.e. attendance and learning) and intermediary 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 midline evaluation report.

The SCW midline 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 in project midline evaluation reports.

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 with both parties providing different perspectives and streams of evidence to generate evidence at the project and window levels. Both the EM, and the majority of projects (see Table 2), collect data in intervention areas and in matched control areas to enable a counterfactual evaluation of the SCW's impact at midline and endline. This means that impact 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 model).

2.1.1 Project 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). For example, Save the Children (Ethiopia) conducts learning tests with girls in grades 1-3, who constitute their primary target group. 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. 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 the refugee camps in Kenya (WUSC project areas). All other SCW projects continued to use a randomised control trial or a quasi-experimental evaluation design (see Table 2 below)[32].

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

	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		EGRA/ EGMA	EGRA/ EGMA	EGRA/ EGMA	EGRA/ EGMA	UWEZO	EGRA/ EGMA		UWEZO	UWEZO	EGRA/ EGMA	National
Learning test conducted in	Mix	Mix		School	HH	HH	School	HH	HH		Mix	Mix	HH	School
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.													

[32] Save the Children (Ethiopia), in consultation with the FM, have stopped tracking a cohort of girls and surveyed a cross-sectional sample of girls at midline. See Section 2.5.2 for more details.

2.1.2 The EM evaluation design

The GEC EM assesses the overall impact of the SCW (at the window level) on girls who were aged 5-15[33] at baseline, and live in the catchment areas of SCW projects' treatment and control schools. The EM conducts longitudinal surveys and learning tests with a random sample of these girls and collects data at three points during the programme's lifecycle. 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's 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 midline, 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 midline, the sample includes 6,084 households and covers both intervention and control areas across the 14 SCW project areas[34].
- A follow-up School Visit Survey (SVS) to confirm the enrolment and attendance rates reported by the girls' caregivers in the household survey. The school visit survey also gathers some contextual information about the girls' schools, classrooms and teachers. The SVS has been conducted in all countries except Afghanistan, Sierra-Leone and Camfed's project areas in Zimbabwe. The SVS covered 2,853 girls at midline.
- In-Depth-Interviews (IDIs) with caregivers, teachers, and community leaders are carried out at baseline and midline to collect nuanced qualitative information about learning conditions in local schools and a range of factors that may prevent local girls from going to school and learning. At midline, researchers probed any changes in the learning environment and asked about any interventions that have taken place since baseline. 744 IDIs were carried out at midline.
- A School-based assessment (SBA) was conducted at baseline and midline in intervention and control schools across SCW project areas in DRC, Ethiopia, and Kenya. The SBA at midline included 406 classroom observations, 427 teacher surveys and 7,237 learning tests with boys and girls that allow us to assess the evolution of gender disparities in literacy and numeracy over the course of the project.
- Review and meta-analysis of the reports, datasets and outcome spread sheets submitted by SCW projects to triangulate the analysis of EM quantitative and qualitative data.

The EM draws on these different sources of evidence to produce baseline, midline, and endline evaluation reports for the SCW, looking at trends and patterns that emerge across the window. A discussion about the differences in EM research activities across different project areas can be found in Section 2.5.1.

2.1.3 Changes to the EM's evaluation design and methodology since baseline

The EM's overarching evaluation design has not changed since baseline, and the fieldwork methodology has only seen minor changes. The midline research was conducted with the same longitudinal sample of girls and their households that was drawn at baseline. New girls and households were only added to the sample to substitute respondents that could not be re-contacted at midline or who refused to be interviewed again. In a few cases, an entire sampling point (defined as a school catchment area, which is typically a village or community) had to be substituted because the one originally surveyed had become inaccessible (details are provided in Annex E).

The research instruments were adjusted to incorporate lessons learned from baseline, which included cutting down the length of the household survey to avoid respondent fatigue (see Annex E for a detailed description of changes to the EM research instruments). An important change to the **EM's qualitative research** approach was that girls did not participate in IDIs at midline. At baseline, many girls had been too shy to share their views and opinions

[33] Or 5-17 in Camfed project areas. See Section 1.3 in Annex E for more details on the EM midline research methodology.
[34] 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).

freely, providing little qualitative information. At midline, girls only participated in the quantitative household surveys.

The **Ebola outbreak in Sierra Leone in 2014** caused Plan International to revise their project intervention strategy, as government schools were closed for nine months. In Sierra Leone, the EM household survey included a special module on Education in Ebola to assess the impact of the Ebola crisis on girls' education.

2.1.4 M&E support provided to projects at midline

The FM supports projects with their evaluation and monitoring systems on a continuous basis. The FM 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 midline:

- Provided projects with a household survey template for midline, as well as a guidance package explaining how the survey has changed between baseline and midline, 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 midline evaluation report template; and
- Reviewed and quality-assured projects' midline research instruments and evaluation reports jointly with the FM.

2.2 Triangulation of different methods and data sources

As described in Section 2.1 our SCW midline evaluation report draws evidence from a range of different data sources to answer the GEC programme evaluation questions. Table 3 presents a simplified version of the GEC evaluation framework, showing how methods and 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					Secondary data
	Project Midline Report	Project Data	Outcome spread-sheet	HH Survey	Learning tests	School Visit Survey	School-Based Assess.	Qual. IDIs	
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's surveys as well as the project datasets and outcome spreadsheets to assess the scale of change in education outcomes and educational barriers that projects across the SCW have achieved since baseline. Projects carry out a bundle 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 data from the qualitative IDIs and the analysis presented in the projects' own midline evaluation reports to unpack how changes have come about and how effective different activities have been at creating these changes.

2.3 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 school-based and community-based interventions would have a sufficient magnitude to be picked up by sampling girls from the local communities. Nevertheless, there is a possibility that some of the girls sampled attend further-away schools that are not receiving any GEC intervention. These caveats are explored and discussed in more detail in Box 2.

The projects and EM samples are independent from each other and are not representative of the same population. 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 the GEC 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 activities. Similarly, control groups could have been dropped from a project sample but not from the EM sample.

In addition, girls in the EM sample might attend a treatment school but not fall within the specific age range targeted by the project, as the EM sampled all girls aged 5-15 at baseline in all project areas[35] for the purpose of consistency and comparability, even though several projects targeted (and sampled) only primary school girls (see Table 6 and Table 7). As a result, **only project data should be used to quantitatively assess the impact and achievement of projects on their target population**. 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, as lack of birth registration is common and families may estimate their children's birth date rather than know it for certain.

In each project area, the EM surveys roughly 400 households and conducts EGRA and EGMA tests with 400 girls that are randomly selected from within the surveyed households. The households and girls are being tracked longitudinally from baseline (in 2013/14) through to midline (2015/16) and endline (in 2016/2017). The sample of 400 households per project area is split relatively evenly between treatment and control school catchment areas. Qualitative IDIs were completed in about 14 per cent of the sampling points chosen for the quantitative sample (see Annex E for further details on our sampling approach).

Box 2: Composition of the EM treatment and control samples in terms of GEC and non-GEC schools, and impact on the findings presented below

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. At the time of writing, eight SCW projects had provided this information.

A quick analysis (described further in Annex F) shows 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 shows that girls attending GEC schools have not performed significantly better in the GEC main outcomes (attendance, literacy and numeracy) than girls attending non-GEC schools[36]. These findings confirm those described further below in Section 3 using the EM treatment and comparison groups.

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

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

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's midline research and mitigation strategies

Survey attrition

The EM household survey and school-based assessment (SBA) are longitudinal studies, which aim to interview the same girls (and boys for the SBA) 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 and SBA, girls and boys that could not be interviewed again at midline were replaced with children from the same school or 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 shared specific characteristics, or if the substitute respondents differ systematically from those who dropped out. Our analysis of attrition from the EM household survey does not show any systematic difference between the baseline and midline samples with regards to key baseline 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 mostly been similar across treatment and control groups, except in AKF (Afg), STC (Eth) and CfBT (Ken). 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. At the window level, attrition rates have been similar across treatment and control areas.

Table 4: EM midline attrition rates by project area

Midline attrition	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 Z-T
Treatment	33%	11%	6%	3%	46%	34%	18%	33%	46%	28%	39%	51%	39%	31%	45%
Control	30%	14%	16%	6%	53%	25%	24%	30%	32%	31%	37%	49%	41%	33%	
p-value	0.93	0.39	0.00	0.21	0.20	0.04	0.18	0.57	0.00	0.49	0.76	0.58	0.68	0.72	

Note: 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.

Box 3: 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 30 per cent at midline in our EM sample. A randomised controlled trial with a 30 per cent attrition 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 midline EM sample (about 6,000 households) is therefore able to detect a 0.2 SD effect at the window level.

However, EM samples are much smaller at the project level which restrains our ability to detect small changes in outcomes. On average, we sampled 400 households per project area. But attrition levels vary widely across contexts. **In project areas with high attrition rates such as IRC (DRC) or RI (Som) we will therefore be less able than in other project areas to capture a significant effect on GEC outcomes.** As an example, under the same assumptions as in the previous paragraph, a sample of 400 allows to detect an effect of about 0.35 SD if attrition is below or equal to 10 per cent. But if attrition is equal to 50 per cent, the same sample will only be able to detect effects larger than 0.45 SD.

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

window level. Any finding expressed at a level of disaggregation lower than window level (such as project level) must be interpreted with those caveats in mind.

Limitations to the analysis of EM data at project level

The 14 SCW projects run different types of intervention in different contexts and with different populations. Their heterogeneity makes it difficult to generalise findings across the window. Ideally, assessing the effectiveness of different activities in different contexts would require disaggregating EM findings at project level. But, as described in the box above, the EM samples are substantially smaller than project samples and were not primarily designed for project-level analysis. 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.

In reality, 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 and midline datasets to the extent that the EM has not been able to draw on project data systematically to unpack the changes observed at the window level.

This means that we can only draw on EM data, in addition to evidence provided in project midline reports, to explore changes between baseline and midline at project level. The small sample 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 and reflected in the EM quantitative analysis at project level.

Spill over effects and contamination

EM quantitative and qualitative analysis shows that education programmes are widespread in SCW treatment and control areas. IDI respondents mentioned GEC activities taking place in areas assigned to the control group and it might be that treatment and control sampling points are too close to each other geographically to prevent direct or indirect 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.

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 quantitative findings with our qualitative analysis and findings from projects.

Differences in EM research activities across different project areas

For practical reasons including safety, projects' own evaluation designs, and the difficulty of getting research permissions, school-based research activities have been dropped in a number of contexts. The EM did not conduct any school-based research (including school visit surveys) in Afghanistan to avoid any tensions that may have arisen in the run up to the national elections at baseline. EM school-based research was also not conducted at midline 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 its project areas.

At baseline, Camfed (Tanzania and Zimbabwe) did not provide a community-based listing as its intervention population was located exclusively within 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 mixture of randomly selected households and purposive sampling of girls who resembled (and included) the target population in terms of receiving bursaries was used (see Section 1.3.1 of Annex B). In the absence of a control group, we cannot calculate a similar difference-in-difference indicator for Camfed as we do for other project areas.

At baseline the EM agreed to conduct additional assessments in schools. Due to budget constraints this was limited to four countries. In Sierra Leone at midline, the school-based assessment (involving classroom observations, interviews with teachers and the testing of girls and boys) and school visit survey were cancelled following the 2014 Ebola outbreak. As a result, the EM school-based assessment (SBA) was conducted in only

three SCW countries at midline, namely DRC (IRC), Ethiopia (ChildHope, STC) and Kenya (WUSC, CfBT), as opposed to four countries as at baseline.

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. We see such differences reflected in the data collected through the qualitative interviews and the EM surveys, especially where respondents were asked about their feelings or aspirations. For instance, 56 per cent of caregivers in Camfed (Tanzania and Zimbabwe) project areas feel there is not enough support in their communities for girls to succeed in school. This proportion is the same as in Save the Children (Ethiopia) project areas although further research showed that community attitudes and aspirations towards education in Afar pastoralist communities is very low compared with Zimbabwe and Tanzania. Our quantitative analysis of changes in barriers between baseline and midline needs to be interpreted with those limitations in mind.

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 may have been tested in a language they speak well in some areas, but not in others. Nevertheless, we score all children against a common metric (i.e. the EGRA oral reading score in words-per-minute) without accounting for such differences.

The analysis presented in this report therefore needs to be interpreted with care and with each project's specific target population and context in mind.

2.5.2 Limitations of the SCW projects' midline research

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

Attrition

As shown in Table 5, a number of SCW projects experienced **high attrition rates at midline.** Save the Children (Ethiopia) experienced the most severe attrition rates. A severe drought that started in 2015 caused communities to leave the Afar region in search of water and grazing opportunities. During a sub-cohort study in June 2015, the external evaluator could only identify 28 of the original 100 girls selected for cohort tracking, despite using GPS and repeated call-backs. In response to this, Save the Children (Ethiopia) abandoned the cohort tracking approach (as agreed with the FM) and adopted a cross-sectional design at midline.

Other project areas where attrition was considerably higher than expected included the Aga Khan Foundation (Afghanistan, 63 per cent), IRC (DRC, 43 per cent), CfBT (Kenya, 48 per cent), WUSC (Kenya, 43 per cent), Camfed (Zimbabwe, 35 per cent), and World Vision (Zimbabwe, 35 per cent). As discussed above, these projects may face greater difficulty in demonstrating significant treatment effects at midline and endline as a result of high attrition.

In ChildHope (Ethiopia) project areas, attrition was not much higher than expected (15 per cent) but it was found to be non-random, meaning that the girls who dropped out of the study were systematically different from the girls who remained in the study, which can introduce a bias to the remaining sample.

Table 5: Attrition rates reported in project midline reports

Midline attrition	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	T	Z
Treatment	30% 12%*	63%	N/R	43%	N/A	15%	43% 18%*	43%	8%	N/R	27% 31%*	12.8 %	23%	18*	35%*

Notes: BRAC (Afghanistan) reported an attrition rate of 30% in government schools, and 12% in community-based schools. IRC (DRC) reported an attrition rate of 43% for *in-school* girls. STC (Ethiopia) are no longer tracking girls longitudinally. Camfed's attrition rates refer to attrition from the reading and maths assessment amongst girls in the treatment group. WUSC (Kenya) reported 43% attrition from the household survey, and 18% attrition from the school-based learning test. Relief International (Somalia) reported an attrition rate of 27% for cohort girls, and of 31% for households. ACTED (Afghanistan) and Plan (Sierra Leone) did not report attrition rates in their reports.

Contextual challenges

Many projects working in fragile and conflict-affected areas faced contextual challenges that impacted on their ability to re-contact baseline households and also disrupted their data collection schedule.

In Afghanistan, ACTED, BRAC and the Aga Khan Foundation struggled with insecurity, instability and displacement affecting regions such as Kunduz (which fell under Taliban control in September 2015), neighbouring Baghlan, Badakhshan and Faryab. Many communities became inaccessible due to fighting, Taliban presence, or displacement. The projects mitigated security risks as far as possible by maintaining flexible data collection schedules and substituting inaccessible sampling points with more accessible ones. However, as reported by the Aga Khan Foundation, this sometimes meant that replacement households had to be selected from already surveyed communities, which interfered with a random sampling approach.

WUSC (Kenya) faced a volatile security situation in some of its project areas in Kenya, due to terrorist threats, including the kidnapping of a teacher in Garissa County. Some teachers left the area and some schools closed temporarily as result, hindering data collection and causing higher attrition rates. WUSC also faced particular challenges with fluctuating populations from South Sudan and Somalia in the Dadaab and Kakuma refugee camps. According to WUSC's midline report about 28 000 new refugees have arrived from South Sudan since 2014, while at least 160 000 Somali refugees were repatriated. This resulted in the creation of new camps and schools, as well as shifts of population between camps and corresponding schools. As a result, cohort tracking posed a significant challenge.

In Kenya, both CfBT and WUSC further struggled with a prolonged teacher strike in 2015 which interfered with the timing and sequencing of midline data collection. Both projects responded by adapting their data collection schedules and prioritising community-based data collection while the schools were closed.

Time constraints

Several SCW projects reported that tight timelines made it difficult to sequence data collection activities in the best possible way. WUSC (Kenya) would have liked to unpack quantitative findings through follow-up qualitative research but were forced into conducting all data collection at the same time due to time constraints. Enumerators for Camfed (Tanzania and Zimbabwe) and World Vision (Zimbabwe) had to cut down on actual survey time because so much time was spent travelling between sampling points. Relief International (Somalia), Plan (Sierra Leone), and CfBT (Kenya) also reported time constraints.

Finding qualified interviewers

In many contexts such as Kenya, Mozambique, Somalia and Afghanistan, projects faced challenges in finding and retaining researchers that had the right research skills and were able to speak local languages and access local communities. Care (Somalia), for example, struggled to find enumerators with the right clan affiliations to access target communities. In particular they struggled to find trained qualitative researchers with the right background, so used the same team for the quantitative and qualitative data collection. The resulting lack of specialised qualitative research skills within the team, combined with cultural factors led to girls providing relatively short responses with limited potential for in-depth qualitative analysis. Projects generally mitigated issues with enumerators' research skills by allowing enumerators to practice with the data collection tools (e.g. tablets or topic guides) and by pairing them with more experienced colleagues and supervisors.

Inaccurate learning assessment scores

Relief International (Somalia) used an adapted UWEZO test to measure numeracy and literacy. UWEZO is designed to assess competencies expected of second-graders, but the project used it to assess fourth and fifth graders. While only 7 per cent of ceiling effects were observed for numeracy scores, ceiling effects for literacy scores increased from 27 per cent at baseline to 48 per cent at midline. This issue alongside other issues with cohort tracking have led the FM to assess RI's learning data as only partly conclusive. Care's (Somalia) data faced similar issues with their literacy scores.

CfBT (Kenya) also had issues with ceiling effects on the UWEZO literacy and numeracy tests at baseline. Additional "extension" tests were given to girls of specific grades to account for these but this made it very challenging to replicate a test with similar difficulty at midline. Numeracy was re-baselined due to issues with extension tests at midline. As a result, the project's learning achievements are only assessed on the basis of literacy scores.

The FM raised questions about the learning scores collected by WUSC's (Kenya) external evaluator. While numeracy scores were relatively consistent from baseline to midline, there was an extraordinary jump in literacy scores. After further inquiry, the external evaluator identified a systematic calculation error during data analysis as the cause of this discrepancy. Learning data was then recalculated to fix this error. However, some concerns about the validity of WUSC's learning data remain, as the project does not have a control group and is not tracking individual girls but uses a panel cohort approach. This means that their external evaluator tests learning periodically amongst a sample of in-school girls who happen to be present in school on the day of the assessment.

Camfed's (Tanzania and Zimbabwe) learning data has been judged partially conclusive by the FM due to concerns regarding how tests were graded as well as a lack of transparency on how aggregate scores were calculated. Camfed uses a national assessment designed by the Ministry of Education to evaluate girls' performance, which makes it difficult to compare scores with any external benchmarks in order to assess girls' trajectories on an absolute scale.

A few projects also faced issues that relate to inconsistencies in the way learning assessments were administered. For instance, some subtasks were timed at midline but were not timed at baseline. Skip patterns also varied across waves and contexts, with some subtasks being given to all students or sometimes to the best-performing only. The method and weightings used for aggregating scores across subtasks could also vary and were systematically reviewed by the FM.

Inability to match midline and baseline observations

Some projects, including Aga Khan Foundation (Afghanistan), Save the Children (Mozambique), and World Vision (Zimbabwe) struggled to match midline observations with the data collected at baseline. This was largely due to inconsistent or inaccurate coding of identifiers for girls and their households, or to baseline identifiers being held by an external evaluator no longer working on the project. This caused delays in the fieldwork as enumerators struggled to re-contact the right girls and households, as well as in data cleaning and processing. It also made the merging of baseline and midline datasets very difficult.

As a result of data discrepancies, Save the Children (Mozambique) were only able to analyse changes in learning scores for a limited cohort sample of 800 girls. Coding errors on various variables, and the external evaluator's inability to match a large number of baseline and midline cases significantly reduces the reliability of the project's findings, and the extent of midline analysis that was possible.

The Aga Khan Foundation (Afghanistan) struggled to match girls surveyed at baseline and midline because some heads of household gave the wrong names for girls and other female household members, or due to uncertainty about the girls' actual ages. In some cases, enumerators claimed to have re-contacted baseline households at midline, but names or ages of girls in the household, or the name of the female carer, did not match the baseline data. In these cases, data was triangulated with other available data to determine whether it was a genuine re-contacted household.

Issues with the control group

Several projects had issues with their control group samples not matching the treatment samples, or being subject to other interventions that might affect their "usability" as controls (e.g. ChildHope Ethiopia).

ChildHope (Ethiopia)'s external evaluator found significant differences between the treatment and control group with regards to educational background and wealth, suggesting that the average socio-economic status was higher in control areas which is typically correlated with better education outcomes. In World Vision (Zimbabwe) project areas, girls in the control group were significantly older, further along in their schooling, and more likely to have books in the household than girls in the treatment group. In addition, there were significant differences in attendance, literacy and numeracy scores that favoured the control group at baseline, which suggests that randomisation was not fully successful.

Quality of school registry data

Virtually all SCW projects struggled with the poor quality or availability of school registry data. Even though many conducted at least one spot check and head count to assess attendance in school, there were often no reliable class registers to compare the head count with. As a result, the majority of reporting on changes in attendance has been inconclusive.

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".[37] DFID's 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."[38] Marginalisation was defined in terms of the education outcomes that DFID wanted the GEC to focus on.

DFID chose not to prescribe the marginalisation factors that projects should focus on to help girls attend school and learn. There was a weak evidence base for the pathways of educational marginalisation and it was not clear which factors would be most critical in each of the different SCW project contexts. The GEC was intended to help strengthen the evidence base by exploring educational marginalisation in a range of different populations and locations. 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 particularly marginalised from education.

Projects responded to this challenge in different ways. ACTED, for example, considered that its target girls were marginalised because they live in the remote, poor and conflict-affected communities of Afghanistan's Faryab province, where female literacy rates are lower than the Afghan average, and cultural norms around the role of women and girls are particularly restrictive. Save the Children (Ethiopia) identified girls from pastoralist communities in Ethiopia's Afar region as marginalised due to the remoteness of their communities and their nomadic lifestyle that made it difficult for them to attend the existing, sedentary schools. ChildHope (Ethiopia), World Vision (Zimbabwe), Camfed (Tanzania and Zimbabwe) and PLAN (Sierra Leone) included socio-economic indices in their household surveys to identify marginalised girls within the samples drawn from their target communities. Camfed, for example, considered orphans, girls affected by disability, and/or girls receiving any form of welfare as *socially* marginalised; and girls who graduated from primary but did not transition to secondary school, or those who at risk of dropping out of secondary school as *educationally* marginalised.

As each project responded to the challenge of finding the most marginalised girls differently, the SCW's target population includes a variety of age groups, grade levels, and population subgroups across a range of different contexts. Table 6 below shows what types of girls each SCW project is targeting, according to project midline evaluation reports:

- **School age:** All SCW projects, with the exception of Camfed (Tanzania and Zimbabwe), are targeting girls in primary school, whereas only four out of 14 projects target girls in upper secondary. Camfed targets exclusively secondary school girls.
- **Educational situation**: All projects are targeting girls enrolled in school. Ten projects also target girls who have never been enrolled, and seven projects target girls who dropped out of school.
- **Social groups:** Eight projects explicitly considered poverty as a criterion for identifying their target girls, even though the communities targeted by the remaining six projects are equally affected by poverty (see Table 6 below). Based on project midline evaluation reports, seven projects target girls with a disability, six target orphans, five target girls from pastoralist communities, four target girls who have been displaced, and two projects explicitly target young mothers.

[37] https://www.gov.uk/girls-education-challenge#girls-education-challenge--the-portfolio-of-projects
[38] DFID (2012), Girls' Education Challenge, Business Case Version 4, June 2012, p. 30.

Table 6: Project targeting – Primary target groups by SCW project

Midline	Projects targeting this group	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
School age															
Lower primary	12	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	
Upper primary	13	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
Lower secondary	9	✓	✓				✓	✓			✓	✓	✓	✓	✓
Upper secondary	4	✓					✓					✓	✓		
Social groups															
Disabled girls	7						✓		✓	✓	✓	✓	✓	✓	
Pastoralist girls	5					✓	✓		✓	✓			✓		
Displaced girls	4							✓		✓		✓	✓		
Orphans	6						✓			✓		✓	✓	✓	✓
Poor/Hunger	8			✓			✓		✓	✓		✓	✓	✓	✓
Young mothers	2								✓			✓			
Other groups[1]	7	✓	✓	✓			✓	✓					✓	✓	
Educational groups															
In-school girls	14	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
OOS girls (never attended)[2]	10	✓	✓	✓	✓	✓	✓		✓			✓	✓	✓	
OOS girls – dropped out[3]	7			✓	✓		✓		✓		✓	✓	✓		

Key

1: "Other groups" include: BRAC, AKF and ACTED: Girls living in remote areas; CARE: Child labourers; and disadvantaged caste / ethnic minority. World Vision: "Marginalisation is the risk of drop out due to food insecurity, being overage, religious affiliation (apostolic or evangelical), low school attendance, family composition/relation to household." WUSC: Girls living in host and refugee communities. ChildHope: Others include child labourers, street children, girls affected by early marriage, and girls affected by risky migration. CfBT: Slum dwellers; STC Mozambique. ACTED: Ethnic and language minorities and remote girls / girls in communities without access to schools.

2: "OOS girls" refers to out-of-school girls (that is girls who are not currently enrolled).

3: "Girls dropped out" refers to girls who were enrolled in the past but de-enrolled prematurely.

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. Table 7 (below) shows how girls from specific target groups are represented in the projects' survey samples, drawing on information from projects' midline evaluation reports. It seems to suggest that girls with a disability only make up a very small share of projects' girl samples, ranging from 3 per cent in Relief International (Somalia) and CARE (Somalia) project areas to 7 per cent in Save the Children (Mozambique) project areas. However, those numbers may not reflect the actual composition of the samples due to the difficulty of capturing survey data on disability and the cultural sensitivity around those issues (see the note at the bottom of Table 7). Likewise, while several projects are targeting pastoralist girls among others, only Save the Children (Ethiopia) works exclusively with pastoralist communities in the Afar region.

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 particular grades at school. Each project sample is representative of the particular target groups that the project is targeting and therefore not exactly comparable to other SCW project samples. The EM's complementary sample is representative of girls aged 5-15 at baseline (and 7-17[39] at midline), living in the catchment areas of intervention and control schools. The EM sample is not specifically tailored to represent the specific target groups of each project to allow for data aggregation and comparisons across the window. As discussed in Annex F these girls sampled by the EM do not necessarily attend a SCW project intervention school (although most do), and are not necessarily targeted by a project intervention, for example because they are not part of the targeted age group. As a result, the EM samples differ from the project samples with regards to their composition.

[39] Or aged 5-17 at baseline, and 7-19 at midline, in Camfed project areas. See Section 1.3 in Annex B for more details on the EM midline research methodology.

Table 7: Representation of target girls in SCW projects' samples

Number or % of total sample	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
Total sample														
All girls	1654	1815	946	4117	…	1,649	…	1454	800	2880	1667	1080	3000	2406
In-school girls	1339	1156	343	3595	…	1519	…	…	428	…	1550	862	…	2406
Out-of-school girls (all)	320	604	603	159	…	80		…		…	66	129	…	
School phase														
Lower primary	47%	97%	36%	55%	…	80%		35%	49%		49%	20%	93%	
Upper primary				45%	…	12%	…	65%	29%		39%	49%	0%	
Lower secondary	52%	3%				8%	…		15%		16%	30%	7%	100%
Upper secondary														
Social groups														
Disabled girls (1)						4%		…	7%	4%	3%	3%	…	
Pastoralist girls					100%	53%		…	42%			11%		
Displaced girls							…		2%		10%	5%		
Orphans						1%			1%		7%	1%	…	40%
Poor/Hunger			…			39%		…	…		22%	100%	…	100%
Young mothers								…			11%			
Educational groups														
In-school girls	81%	64%	36%	87%	…	92%	…	…	54%		93%	80%	…	100%
OOS girls (never attended)	18%	33%	64%	4%	…	5%		…			4%	12%	…	
OOS girls – dropped out	1%	3%		9%		1%		…			0%	8%		

Source: Numbers and percentages have been harvested from the project midline reports.
Note: "…" indicates that the project did not provide any information in their midline report.
Cells are shaded in grey if the group is not one of the project's core target groups.
(1) Please note that the proportion of disabled girls in the projects' may not exactly reflect the actual numbers. Capturing data on disability through household survey is usually challenging due to the different cultural interpretations and understanding of the concept across contexts. Evaluators also faced difficulties in translating of the term to respondents with low literacy levels, which is likely to bias the figures reported.

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

The fact that all projects target in-school girls (among others) and that only one explicitly designed interventions for disabled girls suggests that SCW projects did not systematically or exclusively focus on the *most* marginalised girls within their target communities. The rigorous and systematic targeting of only the most marginalised girls was never a GEC programme requirement[40]. While it is clear that many of the girls targeted by GEC projects can be classified as 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 the wider population that was not targeted. As a result, we cannot assess how marginalised GEC target girls were in relation to the country average or to girls from other communities. It is, however, possible to explore and compare the ways in which girls are marginalised in different SCW contexts. To this end,

Table 8 shows a set of "markers" of marginalisation. These are variables from the EM household survey that indicate the severity of educational marginalisation, as well as of other factors that may hamper girls' educational outcomes.

[40] For a more in-depth discussion of this issue see the EM's GEC Process Review Report (2016), available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/501596/Process-Review-Report.pdf.

- **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 (% of caregivers stating it is difficult to afford schooling for girl), the prevalence of landlessness (% of household that do not own any land of their own), absolute poverty (% of households unable to meet basic needs), and the prevalence of hunger (% of caregivers stating that a member of the family has gone to sleep at night feeling hungry more than 5 days in past year).
- **Markers of "social marginalisation"** include the absence of parents (% of girls living without both parents), disability (% of girls with a disability), and safety and security (% 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 with each project's specific 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 have to pay for school fees and materials, which in turn depends on government policies and/or school practices at the local level.

Table 8: Markers of marginalisation by project area

Markers of marginalisation EM 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 has never attended school)	16%	22%	17%	19%	10%	32%	16%	38%	5%	4%	11%	39%	43%	2%	2%	4%
Poor Literacy (Girl scored zero wpm at ML)	46%	48%	45%	48%	65%	96%	42%	79%	44%	56%	62%	42%	46%	38%	16%	18%
Language difficulties (LOI different from language spoken at home)	49%	13%	7%	50%	86%	76%	6%	97%	92%	93%	95%	17%	8%	51%	33%	38%
Parental literacy (PCG cannot read or write a letter in LOI)	65%	90%	95%	97%	35%	94%	77%	91%	54%	53%	76%	50%	51%	54%	56%	31%
Economic Marginalisation																
Affordability of school (Difficult to afford for girl to go to school)	61%	39%	32%	55%	80%	6%	17%	44%	85%	26%	78%	40%	63%	90%	89%	50%
Land ownership (Household doesn't own land for themselves)	26%	47%	23%	34%	29%	40%	12%	71%	66%	4%	24%	49%	24%	7%	14%	12%
Poverty (Household unable to meet basic needs)	44%	31%	22%	46%	43%	43%	32%	63%	42%	30%	44%	36%	41%	56%	69%	24%
Hunger (Gone to sleep at night hungry more than 5 days in past year)	18%	8%	4%	0%	18%	35%	4%	36%	26%	7%	26%	10%	21%	19%	20%	30%
Social Marginalisation																
Absent parents (Girl lives without her father and mother)	11%	0%	0%	0%	8%	5%	5%	5%	5%	5%	13%	7%	4%	38%	31%	15%
Disability (Girl has disabilities)	9%	0%	1%	1%	16%	0%	6%	10%	17%	10%	9%	8%	6%	11%	19%	9%
Safety and security (Fairly or very unsafe travel to schools in the area)	16%	4%	22%	27%	2%	15%	6%	8%	11%	8%	11%	3%	3%	55%	28%	18%

Table 8 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 and economic disadvantage** (indicated by darker shades of red) in **WUSC (Kenya) and Save the Children (Ethiopia)** project areas. Poor attendance, poor literacy, language difficulties, low parental literacy, and hunger are particularly prevalent in both areas. In addition, WUSC project areas also have the highest rate of landless families, and households that cannot meet their basic needs. 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.** The GEC baseline research suggested that pastoralist girls are particularly affected by a range of educational barriers.
- In all three project areas in **Afghanistan** the large majority of girls (over 90%) have a primary caregiver who is illiterate in the language of instruction, with ACTED and AKF project areas showing higher illiteracy rates than any other SCW project area. 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**, compared with other project areas, **but the affordability of school, poverty and hunger are significant issues, as are social factors** of marginalisation that are less prevalent in other project areas. In both project areas, the percentage of children who live without their parents is much higher than in any other SCW project area (i.e. 38 per cent in World Vision project areas). Our qualitative interviews and thematic research in Zimbabwe indicate that in some cases, adults (including parents and older siblings) migrate to South Africa or Botswana for work and send back remittances to help their children and siblings through school. So for some girls living without parents, it may denote an educational advantage rather than a disadvantage[41]. On the other hand, orphans or children with absent parents living with their relatives may be more likely to provide economic support to the household they live in, particularly if they live with an elderly or if they are too many children in the household. This was not always negatively perceived but seen as a coping mechanism for both the children and their caregivers. There is however some evidence that children living without their parents can be significantly disadvantaged, and at worse, abused. AIDS is likely to be another underlying factor of girls living without their parents. The HIV prevalence rate in Zimbabwe was 15 per cent in 2016 according to the Global AIDS Response Progress Report 2016[42]. As one respondent mentioned:

  "some families will be having many children that are being left behind by deceased parents because people are dying a lot because of HIV/AIDS." (IDI with school staff in Matabeleland, WV Zimbabwe)

[41] Evaluation Manager Girls' Education Challenge Fund (2016). Narrow windows and revolving doors. GEC Thematic Research Report, January.
[42] The report can be found at the following address: http://www.unaids.org/sites/default/files/country/documents/ZWE_narrative_report_2016.pdf

Table 9: Marginalisation factors – 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[43]	48% (2011-12)	85% (2013)	67% (2014)	88% (2008-9)	76% (2011)	78% (2013)	19% (2006)	94% (2014)	83% (2010)
Girls Poor attendance[44]	52%	15%	33%	12%	24%	22%	81%	6%	17%
Girls' Primary Completion Rate[45]	No data	60% (2013)	53.3% (2014)	104.1% (2014)	44.6% (2014)	67.6% (2013)	No data	90% (2013)	77.1% (2013)
Literacy rate 2015[46]	58.15%	86.07%	69.49%	85.90%	76.75%	67.37%	No data	91.75%	87.31%
Economic Marginalisation									
Population in multidimensional poverty[47]	58.8%	72.5%	88.2%	48.2%	70.2%	77.5%	81.8%	28.9%	66.4%

The above table shows statistics at the national level for each of the SCW target countries. It is worth noting however, that the years for the national statistics are not consistent across the countries, and the data for Somalia and Kenya are significantly behind (2006 and 2008-9 respectively) while the EM data from Table 8 was collected in 2015/2016. While the following comparisons are only used for indicative purposes, a couple of interesting findings can be noted. Comparing the prevalence of primary school-aged girls' poor attendance at the national level with the average attendance of girls targeted by WUSC (Kenya), the latter appear substantially more educationally marginalised than the former.

When comparing the national population in multidimensional poverty and the percentage of households expressing that they are unable to meet their basic needs, WUSC again stands out with a target group significantly more economically marginalised than the Kenyan average, while within the same country, CfBT's target group is closer to the national average. Oppositely, girls targeted by World Vision and Camfed Zimbabwe appear to fall significantly behind the national average in terms of economic marginalisation.

While it is not possible to establish that the GEC target girls are the *most* marginalised girls in their communities or regions, the markers shown above do indicate that they **face important disadvantages.** Different combinations of factors prevail in different contexts, creating context-specific and complex drivers of educational marginalisation. 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.

Increased marginalisation due to recent contextual factors

The context and operating environment of some of the projects have been increasingly challenging. This is particularly relevant in terms of safety and security as well as environmental disruptions. In Afghanistan, the United Nations Assistance Mission in Afghanistan (UNAMA) and the United Nations Children's Fund (UNICEF) recorded a significant increase in the number of human rights and protection incidents concerning children in the context of the armed conflict[48]. This was also reflected in our qualitative research. For example, respondents in Baghlan and Kandahar provinces mentioned an acid attack by the Taliban on a girl for attending school and an attack on government facilities. Families in Akhtar also reported experiencing pressure from the Taliban to keep their girls at home – this seems to be particularly intense when girls are of secondary school age.

> *"Older girls cannot go to school because of the Taliban, when families hear acid is sprayed in the face of girls, they are scared and don't allow their daughters to go to school."* (Household Head, Baghlan, AKF Project Area)

Similarly, the situation in CfBT Kenya's targeted communities had also security challenges in terms of community conflict particularly in Turkana. Other environmental challenges have also emerged rendering some of the targeted

[43] 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.
[44] 100% - the national Girls Primary Net Attendance Rate
[45] http://data.worldbank.org/indicator/SE.PRM.CMPT.FE.ZS
[46] http://data.uis.unesco.org/Index.aspx?queryid=166
[47] http://hdr.undp.org/en/composite/MPIchanges
[48] http://unama.unmissions.org/sites/default/files/education_and_healthcare_at_risk.pdf

communities even more marginalised such as the Ebola crisis in Sierra Leone, and severe drought (El-Nino) particularly in Zimbabwe, and in Ethiopia where mass migrations from the Afar region (StC targeted communities) were reported. This also contributes to explaining the relatively high economic marginalisation in those areas.

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

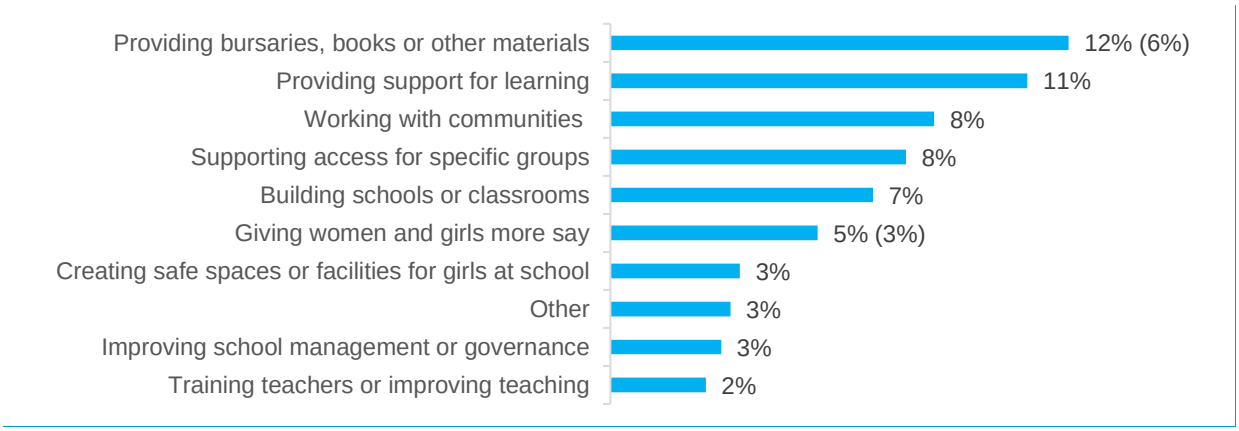
Household survey

A key purpose of the midline evaluation is to understand the extent to which target communities have been exposed to GEC interventions since baseline. The EM household survey asks respondents about changes or activities around girls' education that they may have observed in their communities since baseline, without explicitly mentioning the GEC. This was meant to avoid situations where respondents would be steered toward confirming that a GEC type activity had happened out of politeness.

In treatment areas, 33 per cent of families said that some people or organisations had made it easier for local girls to learn and go to school since baseline. In control areas, 26 per cent of families said the same thing. When asked *who* provided support around girls' education, caregivers in the treatment areas most often mentioned NGOs (20%), followed by local women or women's groups (7%), parent groups (6%), and government officials (6%). The share of caregivers mentioning NGOs working in control areas was significantly lower (14%) than in treatment areas.

When asked about the type of activity that they had seen or heard of, respondents most frequently mentioned the **provision of bursaries, books and learning materials** (see Figure 1). Bursary interventions were mentioned twice as often in treatment areas (12%) than in control areas (6%), which is a statistically significant difference. "Giving women and girls more say" was also mentioned significantly more often in treatment areas (5%) than in control areas (3%), even though the proportion of households mentioning this type of activity appears relatively small, given that promoting the participation of girls in education is a particular aim of the GEC. Other activities mentioned by caregivers were the provision of learning support (11%), work with communities (8%), supporting access for specific groups (8%), and building schools or classrooms (7%). However, none of these were mentioned significantly more often in treatment than in control areas.

Figure 1: Activities mentioned by primary caregivers as promoting girls' education in their area since baseline*



*Note: The figure shows the % of caregivers mentioning these activities taking place in their area since baseline. Percentages shown in brackets show the frequency of mentions in control schools. We only show these if there is a statistically significant difference (at a 5% level) between treatment and control.

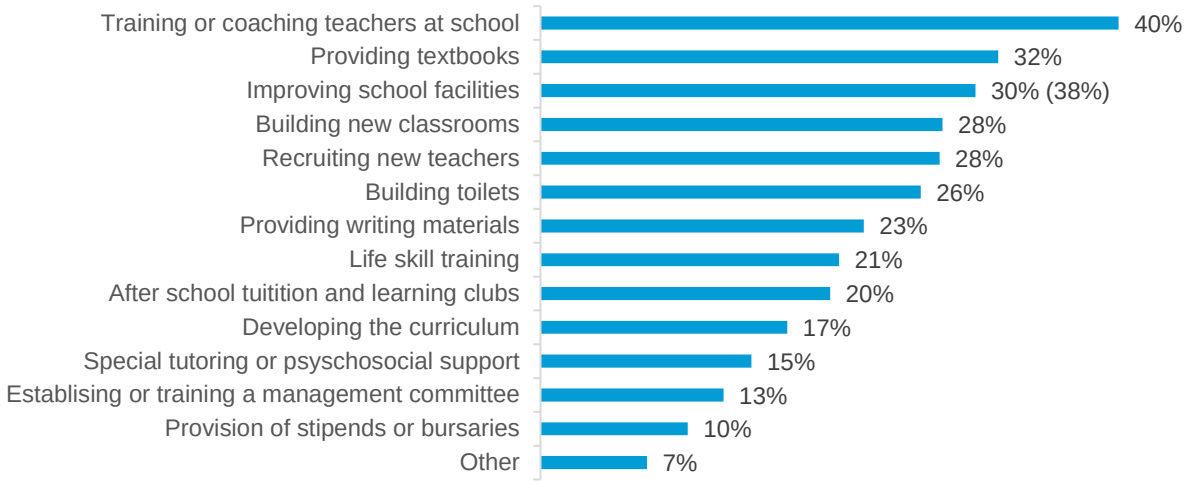
School visit survey

The EM conducted a school visit survey with the school administrators and teachers of all girls who were reported to be enrolled during the household survey. Half of all the treatment schools visited reported that they had been supplied funds by a NGO or religious organisation, compared with only slightly over one third in control areas. There is a significant difference here that might well be attributable to the GEC.

71 per cent of treatment schools reported running ongoing measures to support girls specifically, compared to 62 per cent of control schools. While there is a gap here as well, the share of control schools reporting to support girls' education specifically is still very high. It is difficult to say whether this is due to other actors promoting girls' education in these schools, whether school staff biased their responses to appear more engaged with regards to promoting girls' education, or a mix of the two. We further explore this question in Section 3.1.6.

School administrators mentioned a variety of interventions to support learning that were happening at their schools – but most of these were mentioned as often in control schools as in treatment schools. They included the training and coaching of teachers (40%), providing textbooks (32%), improving school facilities (30%), and many other activities (see Figure 2).

Figure 2: Activities mentioned by school administrators as having taken place in their school since baseline



*Note: The figure shows the % of school administrators in treatment schools stating that this type of activity has happened in their school in the past two years. Percentages shown in brackets show the frequency of mentions in control schools. We only show these if there is a statistically significant difference (at a 5% level) between treatment and control.

The school visit survey also asked each girl's teacher whether the girl had received a bursary in the past two years. According to the teachers, 12 per cent of girls in treatment schools have received a bursary since baseline, compared with only 4 per cent in control schools. This is in line with the information provided by caregivers in the household survey (see above). In addition, teachers reported that 19 per cent of girls have received special tutoring or help, compared with 14 per cent in control schools. Overall, 82 per cent of school administrators in treatment areas said that learning conditions had improved since baseline – but the share was similar in the control group. Similarly, 51 per cent of caregivers in treatment areas thought that learning conditions had improved, compared with 46 per cent in control areas.

While the household and school survey data show a significantly higher exposure of treatment girls to bursary-type interventions and the provision of books and learning materials, other activities (e.g. working with communities, classroom building and refurbishing, creation of safe spaces, teacher training) were mentioned to the same extent in treatment and control areas. In summary, the data suggest that there has been a lot of activity related to girls' education in SCW project areas, some of which may have been driven by non-GEC actors.

3.1.4 Who has been exposed to GEC interventions?

This section presents a set of graphs showing the extent to which different groups of girls have been exposed to GEC interventions, by subgroup and by project areas. **Most subgroups and project areas do not show any statistical difference between the treatment and control groups for any of the variables presented**. **This was expected as our EM sample is not powerful enough to capture changes at lower levels of disaggregation than the window level** (see Section 2.5). As a result, charts are presented for treatment groups only. In the only case where the difference-in-difference indicator is statistically different from zero at the 10% level, we added an asterisk next to the midline treatment number and described values for the control group in the narrative. When the background of the midline number is green, the DID is positive. When it is orange, the DID is negative.

Figure 3 shows that many more girls in secondary school were receiving bursaries, than girls in primary school (37% compared to 12% in lower primary school, at midline). This is likely to be linked with the fact that primary school is free in most GEC countries[49], whereas attendance in secondary school is subject to fees. The share of

[49] 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.

bursary recipients is also above average among girls living without their parents (32%), girls with a disability (24%) and girls whose families struggle to afford the cost of schooling (22%). This suggests that projects running bursary schemes for particularly vulnerable girls, such as Camfed (Tanzania and Zimbabwe), successfully prioritised orphans, disabled girls, and girls living on welfare for participation in their schemes (see Section 3.1.1).

At the same time, the share of bursary girls has dropped slightly since baseline among secondary school girls, and increased only relatively little for girls living without their parents. This may suggest that bursaries are (also) being provided by non-GEC actors to these particular target groups, or reflect the fact that organisations like Camfed (Tanzania and Zimbabwe) were already providing bursaries in target communities before obtaining GEC funding. Across most other subgroups, the share of bursary girls has increased by roughly 10 percentage points since baseline. However, this increase is not statistically different from the increases observed in control areas.

Figure 3: % girls reported as having received a bursary by their caregiver, in treatment areas

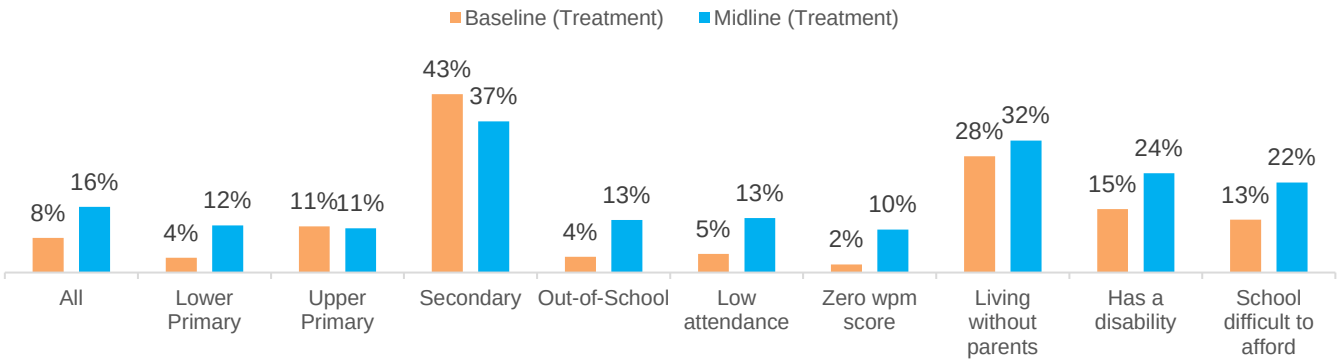


Figure 4 below shows the share of girls who received school books or learning materials at baseline (6%) and midline (10%). The share increased in particular for girls in primary school and virtually remained the same for girls in secondary school. There was also a relatively large increase among out-of-school girls and girls with poor attendance, although it is not statistically different from the increase observed in control areas.

Figure 4: % girls who have been given school books according to their caregiver, in treatment areas

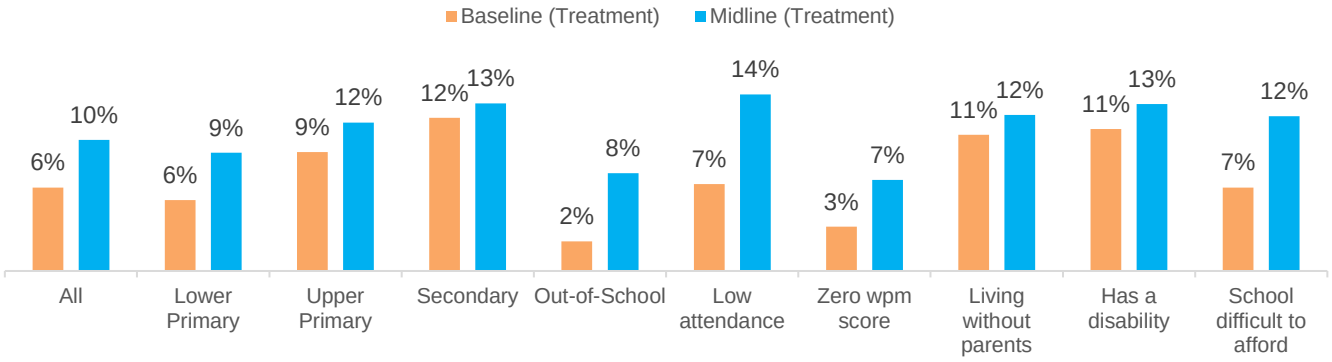


Figure 5 below shows the share of girls who attended special classes or study groups in the year(s) preceding the interview. On average across all groups, there was virtually no change from baseline to midline with roughly 10 per cent of girls having attended such classes at both points in time. The only group that increased their participation in special classes or study groups substantially (by 11 percentage points) were girls with a disability, but control areas show a similar pattern. The share of upper primary and secondary school girls attending special classes or study groups dropped by four percentage points since baseline (from 18% to 14% for upper primary and 16% to 12% for secondary girls). This suggests that SCW projects have tended to prioritise early grades and specific marginalisation groups for special classes or study groups, rather than girls who are further ahead in school.

Figure 5: % girls reported as having attended special classes or study groups, in treatment areas

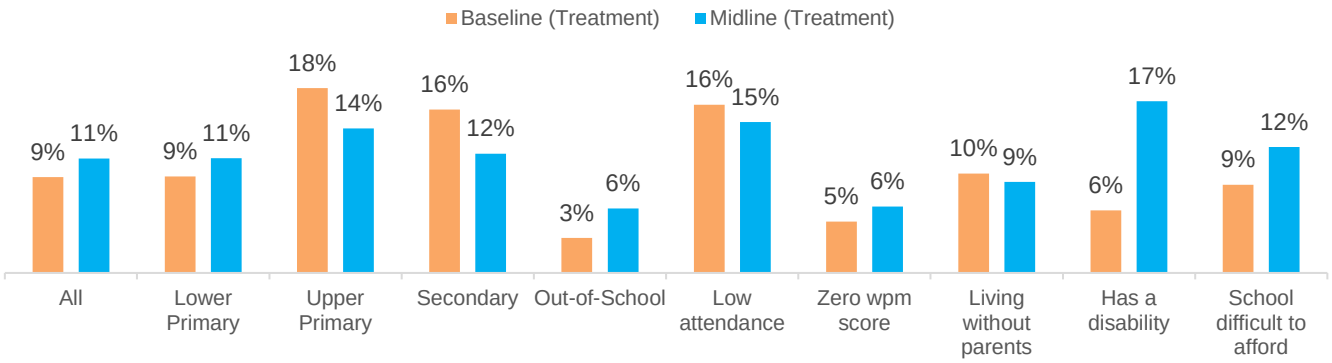
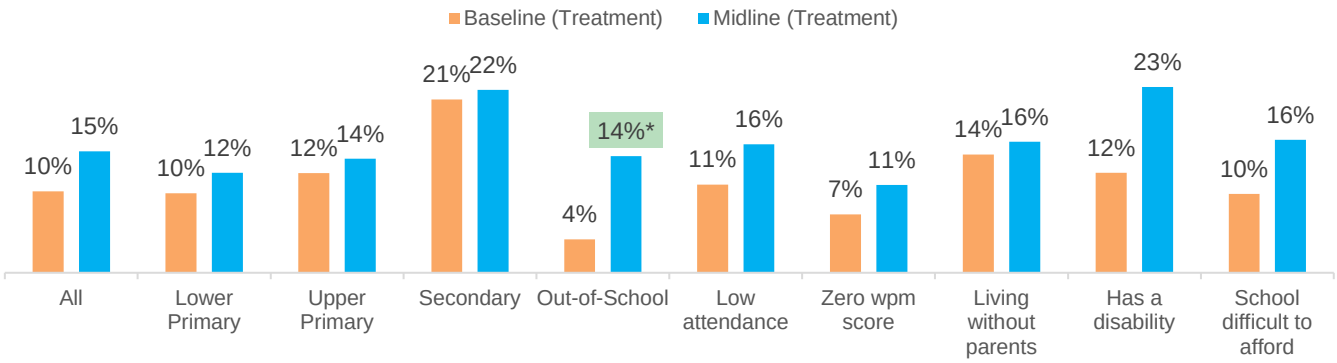


Figure 6 shows the share of caregivers who reported that their girls may have been reached by efforts to enrol them or help them remain in school. At midline, 15 per cent of caregivers reported that somebody had spoken to their girl about enrolling or staying in school, compared with 10 per cent at baseline. The data suggests that out-of-school girls and girls with a disability have been particularly exposed to such efforts since baseline with an increase of 10 percentage points. For out-of-school girls, the difference-in-difference is statistically significant: The share of those reached by enrolment efforts rose from 4 per cent at baseline to 14 per cent at midline in the treatment group, while reaching only 8 per cent in the control group.

Figure 6: % girls who have been contacted about enrolling or staying in school according to their caregiver, in treatment areas



In summary, this analysis reflects the predominant focus on primary school girls in the GEC SCW, with little to no increase in exposure to the above-mentioned interventions among girls enrolled in secondary school. It also seems that out-of-school girls and girls with a disability became more exposed to different types of interventions between baseline and midline. Generally, however, exposure to GEC-type interventions increased both in the treatment and control areas and **there is no significant net effect on treatment girls with the exception of out-of-school girls who have become more exposed to efforts to enrol and keep them in school**, than girls in the control group.

3.1.5 Exposure by project area

We ran the same calculations as in the previous section while disaggregating exposure variable by project area instead of school phase and socio-economic subgroups. Except in one or two cases, we show no statistically significant difference between the treatment and control areas across any of the SCW projects. In some project areas, the reported exposure to certain activities has largely improved from baseline to midline. But it has usually improved in similar proportions in control areas. Besides, the activities described by respondents do not systematically align with the activities that projects reported conducting under the GEC in their target areas. The corresponding graphs and narrative can be found in Annex H.

This analysis first shows that survey-based measures of exposure and reach are subject to important biases, including subjectivity and recall biases. As a result, reported measures, whether from household members (here the girl's primary caregiver) or school personnel (the girl's school administrator or head teacher), cannot be

interpreted literally and do not necessarily reflect the type and scope of the activities conducted on the field. Besides, it is worth reminding that the EM sample is designed to be representative of target communities, not of projects' direct beneficiaries. This implies that such an analysis cannot be used to draw any conclusions on the extent to which projects' activities effectively reached their target population.

The second learning point is that various education-related activities, whether taking place under or not under the GEC, seem to have happened in both treatment and control areas between baseline and midline. GEC-related activities may have taken place in control areas, and oppositely, non-GEC-related activities may have taken place in treatment areas. The next section further explores and analyses those patterns in the light of qualitative data and shows that projects' areas are a crowded space for education interventions.

The previous point implies that the ability of our sampling design to detect a significant treatment effect and to attribute it to the GEC may be substantially reduced. In particular, it is very unlikely that our quantitative sample will detect any impact at the project level. If it does, this impact may not be entirely attributable to the GEC activities. Those caveats need to be kept in mind throughout the report when interpreting project-level findings from EM data.

3.1.6 Who else is working in the same space as the GEC?

Are other organisations working in GEC treatment and control areas?

To better understand who is working in SCW target (and control) communities the EM asked respondents in the qualitative interviews to describe any interventions around girls' education that they had seen or heard of. The evidence collected in this way is summarised in Table 10 and Table 11 below. It is important to note that as mentioned in Section 2 and described in Annex F, GEC projects themselves often ended up operating in areas they initially flagged as control.

Table 10: Potential non-GEC activities taking place in treatment and control areas

EM midline qualitative data	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 AREAS														
Economic interventions	A		B	A	B	A	A	A	B	A	A	A	A	A
Infrastructure and resources	A	B	A	A	A	A	A	A	A	A	A	A	A	A
Teacher training and support	A	A	A	A	A		A		A	A	A	B		
Community-based activities			B	A		A	A	A	A	A		B	B	
Extra-Curricular Activities		B	B	B				A	A	A	B	B	B	
School management & Governance													B	
Empowerment & Self Esteem									B	B				
Marginalisation									A					B
Violence against Children														
CONTROL AREAS														
Economic interventions		A	A	A	A	B	A	A			A	A	A	
Infrastructure and resources		A	A	A	A	A	B	A	A		A	B	B	
Teacher training and support	A	A	A				B			B	B		A	
Community-based activities	B	B	B	B		B	A	A		A	B			
Extra-Curricular Activities		A							B		B			
School management & Governance														
Empowerment & Self Esteem						B						B		
Marginalisation						A		B	B		A			
Violence against Children								A						

Key		
	A	Strong Evidence: many respondents have noted the activities; or the respondent is a school staff; or many actors are reported to carry out this activity.
	B	Weak evidence: possible past activity; only one or two respondents excluding school staff reported on this activity.

Note: This table is based on the analysis of qualitative interviews with caregivers, school staff and community leaders. The table indicates whether respondents in each project area mentioned other actors than a GEC project (or project partner) conducting activities around girls' education, and if so, in what area of intervention.

As shown in Table 10, the qualitative data confirms that SCW communities are a crowded space for education programming. The evidence suggests that non-GEC activities are happening in most treatment and control areas and address similar educational barriers:

- In SCW treatment areas, non-GEC actors have been reported to improve school infrastructure and resources in 13 areas, conduct economic interventions in 12 areas, and run teacher training in 9 areas.
- In control areas, improvements to school infrastructure have been reported in 11 areas, economic interventions reported in 10 areas, and teacher training in six areas.

Several projects mentioned issues with other organisations working in their SCW target communities. WUSC (Kenya) mention a number of international NGOs and relief agencies such as Islamic Relief, Oxfam, UNICEF, Danish Refugee Council and the Red Cross, who all support the construction of toilets, classrooms and water points in Dadaab and Kakuma refugee camps. World Vision (Zimbabwe) reported that respondents in the qualitative interviews were not always able to identify whether an intervention was run by World Vision or another organisation, because of the large number of organisations working in the same space. Another example was provided in the CfBT midline evaluation report (see Box 4).

Box 4: Example of a crowded programme environment

In their midline evaluation report, CfBT reported a lack of governance and coordination at the local and regional level, which reportedly led to the overlap of different education projects with diverging approaches and methodologies in the same schools. For instance, CfBT stressed that in Kenya's Coastal Region, the Aga Khan Foundation (AKF) expanded one of their education projects to 27 CfBT treatment schools without previously consulting the CfBT team. CfBT argue that there is "no synergy" between the two projects and that AKF is using a methodology to teach literacy that is "the exact opposite" of that used by CfBT. According to CfBT's project midline report, this has caused "tremendous confusion and the efforts to harmonise and coordinate are yet to yield results".

Who is working in GEC *control* areas?

The fact that GEC-type interventions are reportedly undertaken by non-GEC actors in a majority of control areas might explain why we observe no significant difference in the exposure of GEC respondents, to different types of interventions. Table 111 shows *who* was said to intervene in SCW control areas.

Table 11: Overview of actors reported implementing (girls') education activities in control areas

EM midline qualitative data	All	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
GEC	11	A	A	A	A			A	A	A	A	A	A	A	
Government	9	B	A	A					A	B		A	B	A	B
International Organisations															
UNICEF	7		A		A			B		B		A		A	A
World Food Programme	2					A							A		
INGOs															
World Vision	4		A					B		A		B			
Save the Children	2			B								A			
Camfed	2											B		A	
Mercy Corps	1											B			
CARE	1											A			
Concern the World	1											A			
Action Aid	1								A						
Oxfam	1							A							
IRC	1							A							
International Institute for Rural Reconstruction	1								A						
Norwegian Refugee Council	1			A											

Development Agencies															
Japanese Aid	1						A								
Korean Aid	1						A								
USAID	2				B	A									
Other NGOs and Development Agencies	12		A	A	A	A	B	A	A		A	A	B	A	A

Key		
	A	Strong Evidence: many respondents have noted the activities; or the respondent is a school staff; or many actors are reported to carry out this activity.
	B	Weak evidence: possible past activity; only one or two respondents excluding school staff reported on this activity.

In 11 SCW *control* areas, respondents had seen or heard of GEC project activities that were ongoing in the respective project target communities. This confirms that treatment and control areas were not accurately defined from the outset (again, see Annex F for more details). Besides, activities and / or changes may have spilled over from these treatment areas. Household survey data shows that in some cases control and treatment sampling points are located very close to each other, and might even be neighbouring villages, which may have favoured spill overs from treatment to control communities.

The qualitative interviews suggest that some spill over effects could be the result of control communities mobilising themselves to replicate an activity carried out in a nearby GEC treatment area. This was also reported in some of the project midline evaluation reports. For example, Save the Children (Mozambique) report that girls who changed schools established girls' clubs at their new school, and that members of a women's group in a control community independently sought literacy classes for themselves. While such spill over effects make it more difficult to capture the GEC's effect statistically they are a positive sign of impact as they show that communities recognise the benefit of certain GEC activities.

Other actors that were mentioned as carrying out education interventions in control areas included government (mentioned by respondents in eight projects), UNICEF (five project areas), World Vision (three project areas), Save the Children and the World Food Programme (both in two project areas).

These examples underscore the notion that GEC project areas are indeed a crowded space. They also show that exposure to GEC activities is difficult to measure based on family and community perceptions, as respondents may not be aware of all the activities that are ongoing, and of the organisations running them. Respondents are also more likely to reference tangible activities that they have directly received, seen or heard of (like material support).

Summary of key findings – Relevance: To what extent has the GEC reached marginalised girls?

Our analysis of marginalisation shows that SCW target **girls are disadvantaged across a variety of dimensions that differ across project areas**. For instance, WUSC (Kenya) and Save the Children (Ethiopia) project areas show the highest concentration of educational and economic disadvantage due to the particular contexts that they are working in. Conversely, Camfed (Tanzania and Zimbabwe) and World Vision (Zimbabwe) project areas show relatively low levels of educational marginalisation but poverty, hunger and social markers of marginalisation are higher than elsewhere.

Data on exposure to GEC-type interventions suggests that **primary school girls were reached in larger numbers than secondary school girls**. This is in line with the fact that 13 out of 14 projects are targeting primary girls, while only nine and four projects target lower and upper secondary school girls respectively. The evidence further suggests that bursaries have been largely targeted towards secondary school girls, which reflects the fact that primary school is free in all SCW countries (although, in practice, primary students are still required to pay fees in some countries). But the share of girls receiving bursaries did not increase significantly more in the treatment than in the control group between baseline and midline. The data was equally inconclusive with regards to other activities such as the provision of school materials, and study groups. However, **significantly more out-of-school girls were reportedly reached by efforts to enrol them in school** than in the control group.

In general, the proportion of households that report exposure to GEC-type activities was greater at midline than at baseline, but **levels of exposure to education support were similar in treatment and control areas**. Among all education-related activities, caregivers most frequently mentioned girls receiving bursaries, followed

by books and learning materials. Community-based and school-based activities were less frequently mentioned by caregivers which may reflect a lack of awareness of such interventions at the household level rather than an actual lack of activity. Respondents may be biased towards reporting material or financial support (e.g. bursaries or provision of textbooks) which is easier to notice and remember. Respondents may also be biased by the extent and quality of a project's communication of their activities.

School administrators most frequently reported activities relating to the training or coaching of teachers at school, the provision of textbooks and the improvements of school facilities and classrooms. Provision of bursaries, which has generally been delivered directly through households, was less often mentioned in schools, relative to other types of activities, which shows that **levels of perception and awareness can vary across different types of respondents in the same area**.

Most activities (e.g. working with communities, classroom building and refurbishing, creation of safe spaces, and teacher training) were mentioned to the same extent in both treatment and control areas. This may be explained by the fact that **many GEC and non-GEC activities are ongoing in GEC treatment and control areas,** as suggested by our qualitative research. Besides, an analysis of the EM sample across five project areas shows that a substantial proportion of girls from control areas are attending GEC schools (see Annex F). These "spill over" and "contamination" effects make it more difficult to capture the effect of the GEC statistically, both in terms of reach (as captured through exposure questions) and impact (on girls going to school and learning).

Key lessons learned

- While we cannot assert whether girls targeted by SCW projects are the most marginalised girls in their communities or regions, they are disadvantaged to varying degrees and across different dimensions. **Economic and social markers of marginalisation differ substantially from one context to another,** underscoring the diverse range of environments, populations and challenges that SCW projects are working with. Pathways into educational marginalisation are likely to vary, depending on the interplay of different marginalisation factors in each local context. This implies that activities aiming to improve girls' learning have to be context-specific and tailored to girls' specific needs. This highlights the importance of **analysing key 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 analysis of markers of marginalisation suggests that there are **two layers to this marginalisation: the first layer affects all children in project areas regardless of gender, whereas the second layer entails more specific barriers for girls in particular.** Our baseline research showed that the most important barriers to girls' education were families' inability to afford the cost of school, as well as a lack of teacher training and a lack of school equipment and infrastructure. These factors are likely to affect both boys and girls. In addition, however, families often prioritise boys over girls where resources are scarce, and boys may cope better with malfunctioning toilets and poor sanitation in schools. To fully understand and distinguish the universal and gendered effects of these marginalisation factors on education, **gender analysis should be an explicit part of projects' design and M&E processes**.
- SCW communities are a crowded space for education programming and our evidence suggests that **GEC and non-GEC activities are taking place in most treatment and control areas and address 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 how difficult it is to implement valid quasi-experimental evaluation designs in the GEC context where actors, target populations and intervention designs keep shifting.
- A key learning point is that **quasi-experimental evaluation designs require strong evaluation expertise on the ground, as well as extensive quality assurance on the part of the EM and the FM**. The lack of detailed local maps showing where communities are located makes it particularly difficult for the EM team to ensure that treatment and control areas are sufficiently far apart from each other. In future, selecting a smaller but carefully selected and matched control group, and obtaining detailed

information on activities that are ongoing in these areas can help to mitigate some of the challenges of conducting quasi-experimental evaluations in similar contexts.

- Given how many different actors work in GEC treatment and control areas, there may have been **missed opportunities for SCW projects to coordinate with these actors**. Projects were asked to evidence and report in their sustainability sections instances where they had coordinated with other actors in their project environments to deliver complementary benefits. Most project midline evaluation reports though provided little substantial evidence of SCW projects considering potential complementarity between GEC and non-GEC agencies and programmes. In such crowded spaces, it is critical that projects carry out a mapping exercise of the intervention landscape as an integral part of their design and baseline research processes to ensure complementarity as well as avoid duplication and inefficiency.

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

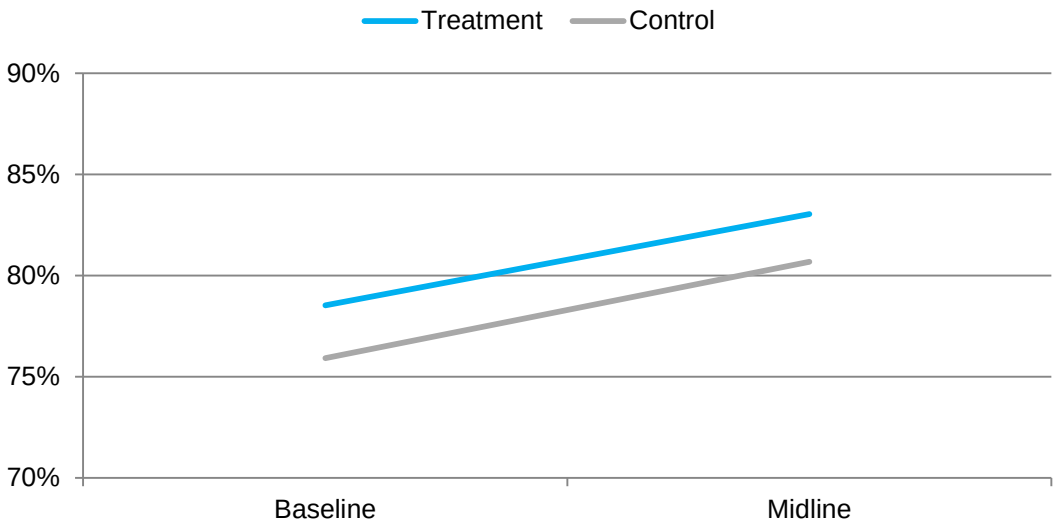
This section, including the analysis of gender gaps in enrolment and attendance rates, relies exclusively on EM data collected through the baseline and midline household and school-visit surveys. Figures showing baseline to midline changes at the window and project level are based on cohort girls only. These are the girls who have been successfully tracked from baseline to midline. Other graphs and figures, except when mentioned in their title, are based on the full sample including girls lost after baseline and girls substituted at midline. Asterisks in bar charts show a significant difference-in-difference at the 10% level across the treatment and control groups between baseline and midline. When the background of the midline figure is green, the DID is positive. When it is orange, the DID is negative.

Camfed project areas do not have a control group in EM data. The project did not provide a community-based listing of treatment and control groups at baseline as its intervention population was located within schools. The EM therefore asked for a list of the home communities of girls due to receive bursaries through the project intervention and used it as a sampling frame. In these communities, a mixture of randomly selected households and purposive sampling of girls who resembled (and included) the target population in terms of receiving bursaries was used. As a result, the whole sample is considered as part of the treatment group. More details are given in Annex E.

3.2.1 What effects has the SCW had on enrolment?

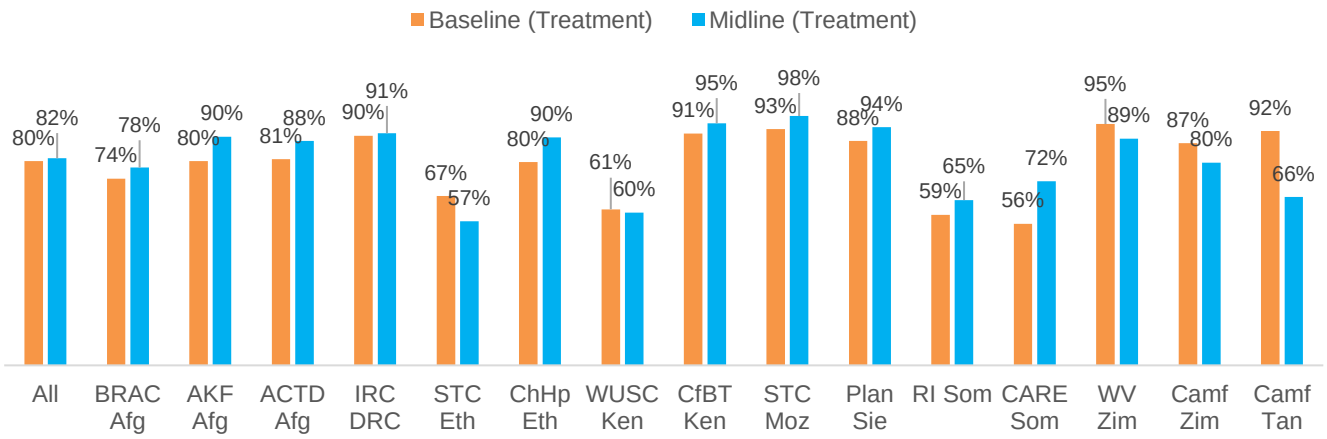
The enrolment rate increased by 4 per cent in the EM's SCW treatment cohort between baseline and midline, from 79 per cent to 83 per cent (see Figure 7). In the control cohort, enrolment increased roughly to the same extent. There is no significant difference-in-difference between the two groups. We therefore cannot ascertain an effect on enrolment at the window level.

Figure 7: Enrolment rates across SCW EM cohort excluding Camfed areas



At project level, no project shows any significant effect on enrolment rates for the treatment cohort compared with the control cohort. Focusing on the treatment cohort only (see Figure 8), enrolment increased in most project areas, with the exception of Save the Children (Ethiopia), WUSC (Kenya), World Vision (Zimbabwe), and Camfed project areas (Tanzania and Zimbabwe) where enrolment rates fell between baseline and midline.

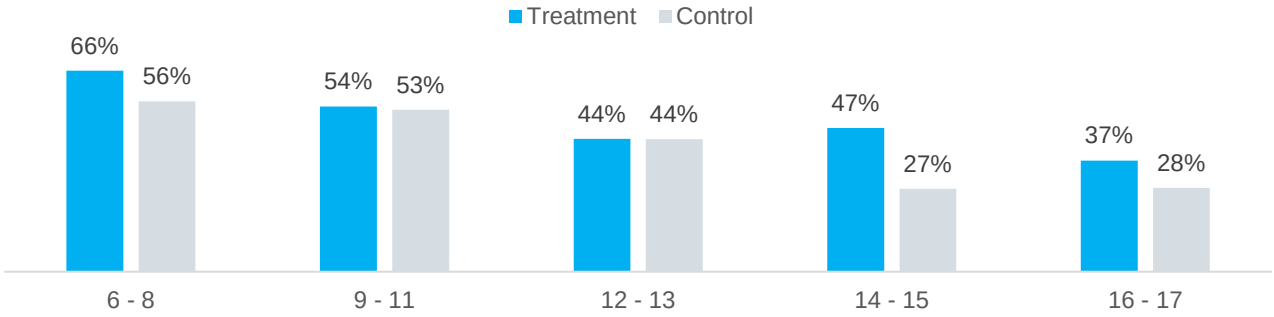
Figure 8: Enrolment rates by project area (treatment only), cohort including Camfed areas



The increase in enrolment measured in most project areas is primarily driven by a rise in enrolment among younger girls (aged 6-8 at midline). These girls were only 5 years old and not yet at school when they were first surveyed at baseline. Among the girls aged 6-8 at midline, enrolment rose from 69 per cent at baseline to 78 per cent at midline. There was virtually no change in enrolment amongst 9-11 year olds, a small increase amongst 12-13 year olds (5%-points), and a small decrease amongst 14-15 year olds (-2%-points). The decrease in enrolment observed in Camfed areas can be explained by the older age of girls targeted by this project (aged 15 or older at midline) who are more likely to drop out-of-school as they reach secondary-school age, or to finish school for the oldest.

Figure 9 shows the percentage of girls in the treatment and control group who were out of school at baseline, and had enrolled by midline, excluding Camfed project areas. Across all age groups, more out-of-school girls enrolled in treatment than in control areas. The difference is particularly pronounced among 6-8 year olds, and among older girls (aged 14-17).

Figure 9: % girls out of school at baseline who now are enrolled at midline



At the same time, the proportion of girls enrolled at baseline who are out-of-school at midline is similar across treatment and control areas (see Figure 10 below). This suggests that the SCW may have had a greater effect on enrolling out-of-school girls than on retaining enrolled girls in school. However, the effect is not significant at window level.

Figure 10: % girls enrolled at baseline who are no longer enrolled at midline, by age category

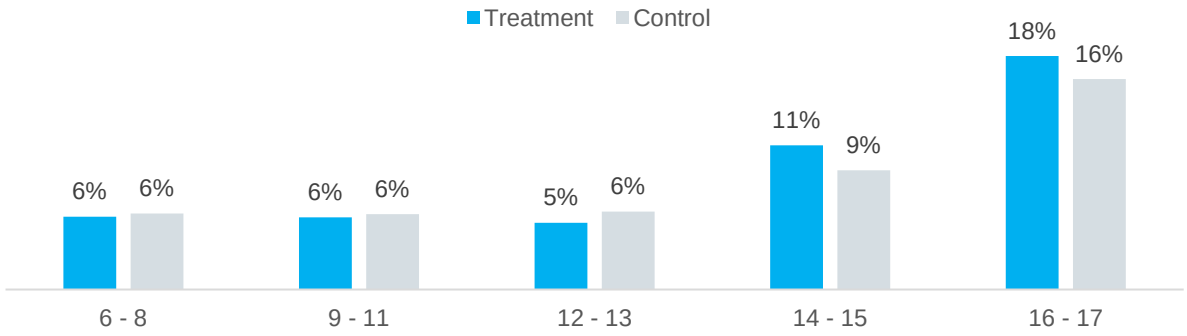
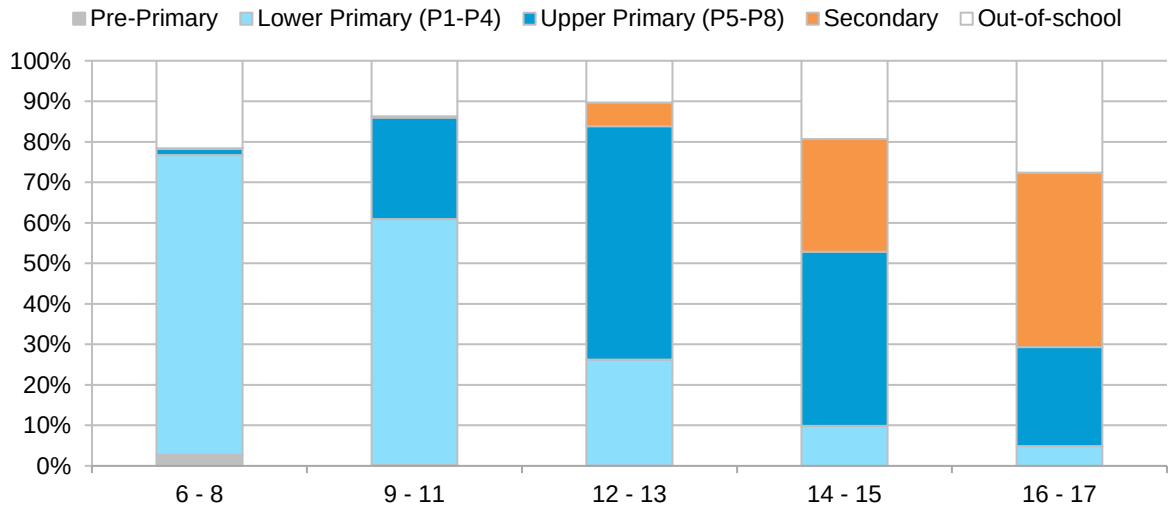


Figure 11 breaks down the sample by the school phase that girls were enrolled in at midline, in relation to their age. As could be expected, the large majority of 6-8 year olds are enrolled in lower primary school, with about 20 per cent not (yet) enrolled, and a small fraction either in nursery, or in upper primary. At ages 12-13, the share of girls enrolled in secondary school is still very small around 5 per cent. Even at ages 14-15 (official secondary age) half the girls in the EM sample are still in primary school. At ages 16-17, this is still the case for 30 per cent of girls. The proportion of out-of-school girls is the smallest at ages 12-13 and the highest at ages 16-17. This confirms the baseline finding that most girls lag behind their schooling age, as virtually all girls from aged 12 should be attending secondary school. Besides, girls start dropping out-of-school at an early age, before the age of 16-17.

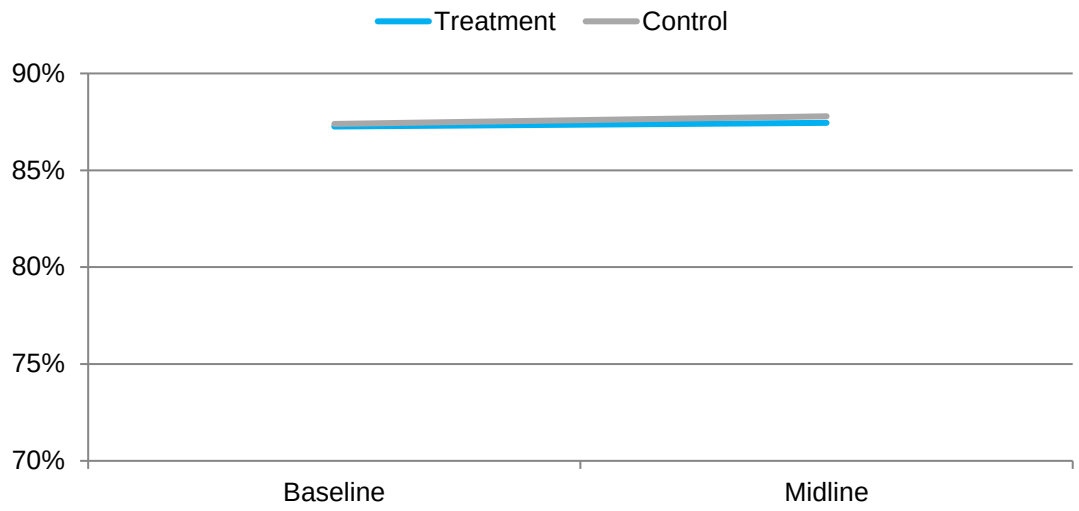
Figure 11: Midline girls at different school phases, by age group (treatment and control areas)



3.2.2 What effects has the GEC had on attendance?

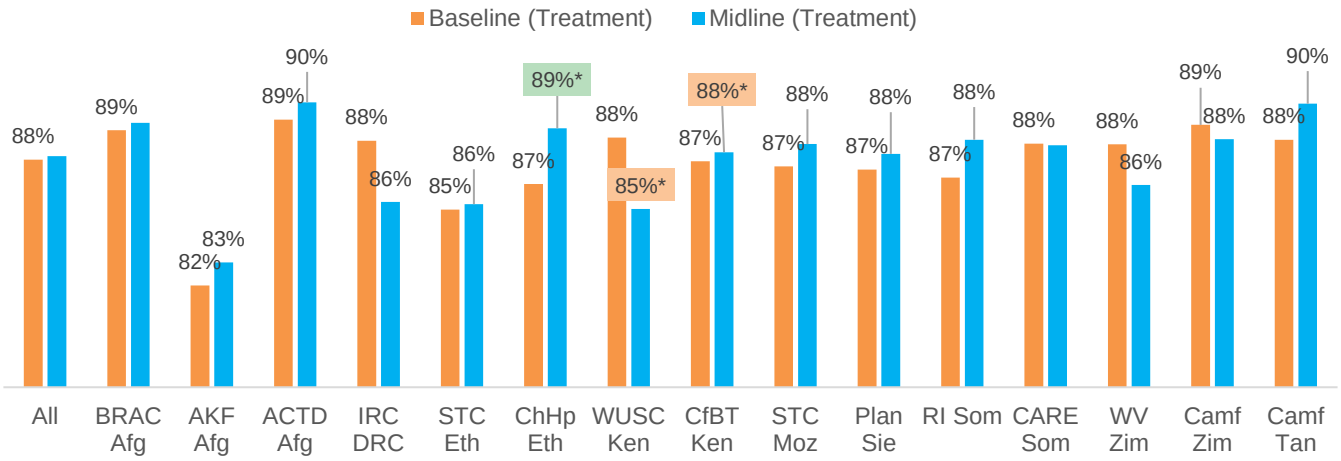
Improving attendance is one of the GEC's key objectives. Attendance is defined as being present in school, as opposed to being enrolled (but not necessarily in regular attendance). Through the EM household survey we measure virtually no change in attendance in the treatment and control groups at the SCW level. As shown in Figure 12 attendance rates remained stable at around 87 per cent.

Figure 12: Attendance rates across SCW EM cohort excluding Camfed areas



At project level, attendance largely remained at baseline level, or increased only marginally (see Figure 13). Two projects, WUSC and CfBT (Kenya) show a significant negative difference-in-difference, while one project, ChildHope (Ethiopia) shows a significant and positive effect on attendance. But none of all SCW project areas exhibit a difference larger than three percentage points between baseline and midline.

Figure 13: Attendance rates by project area (treatment only), cohort including Camfed areas

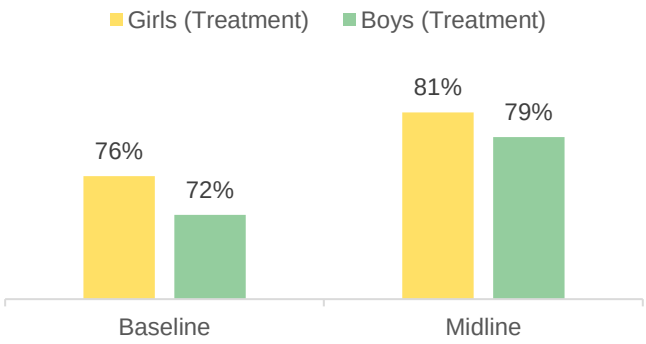


It is likely that the variables measuring attendance in the household survey are not sufficiently sensitive to pick up small changes in attendance. At the same time, projects have equally struggled to measure attendance reliably, using spot checks and school registry data. Poorly kept student records hamper the reliability of both approaches. At midline, the EM included a refined set of variables measuring attendance in the school visit survey. While we lack a baseline comparison, we hope to be able to provide a more reliable midline to endline comparison at the end of the GEC. At midline, the school visit survey (SVS) data suggests that girls attended 89 per cent of school days in the running school year in the intervention group, and 87 per cent of school days in the control group which is broadly consistent with the household-based measure.

3.2.3 What effects has the GEC had on differences in enrolment and attendance between boys and girls?

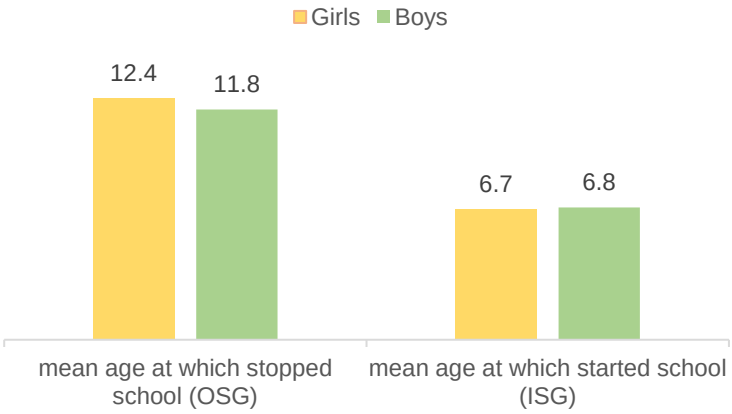
On average across the SCW, the EM data shows **lower enrolment rates for boys than girls in treatment areas**, both at baseline and at midline (see Figure 14). The gap between boys and girls narrowed slightly, with the boys' enrolment rate increasing by 6%-points, compared with 4%-points among girls. This data underscores the EM's finding at baseline that girls are not systematically disadvantaged compared to boys when it comes to enrolment in the SCW project areas. Control groups show the same pattern of change and we don't observe any statistical significant difference between treatment and control areas for either boys or girls.

Figure 14: Enrolment rates for girls and boys – average across SCW



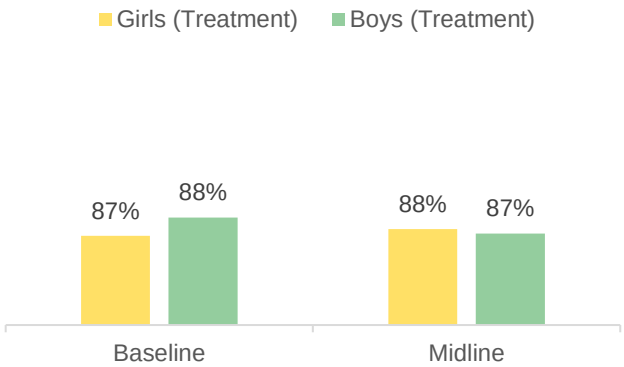
Although one could argue that boys' enrolment has increased more than girls' enrolment due to girls entering school later, Figure 15 shows that across the SCW both girls and boys tend to enter school around age 7. Out-of-school children have also left school at around the same age, with girls being slightly older upon dropping out (12.4 years compared to 11.8 years for boys).

Figure 15: Ages at which boys and girls started and stopped school - average across SCW



We measure **equal attendance rates for boys as for girls** based on household survey data, as shown in Figure 16 below. Attendance for boys has remained stable since baseline, with only a slight decrease among boys in treatment areas (88% to 87%). As mentioned above, the lack of variation in attendance may be due to the EM household survey not picking up subtle variations that exist in the population.

Figure 16: Attendance rates for girls and boys – average across SCW



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

This section relies exclusively on EM baseline and midline data, with the exception of the final comparative analysis of EM and projects' findings regarding literacy and numeracy outcomes which is based on both EM and projects' data. The gender gap learning analysis is based on data from the EM's school-based assessment that was conducted in five project areas across three countries (IRC in DRC, Save the Children and ChildHope in Ethiopia, WUSC and CfBT in Kenya). The rest of the analysis relies on EM household and school-visit survey data. More details on the EM's survey methodology are provided in Annex E. As in the previous section, figures showing baseline to midline changes at the window and project level are based on cohort girls only. Other graphs and figures, except when explicitly mentioned in their title, are based on the full sample including girls lost after baseline and girls substituted at midline.

As far as possible, difference-in-difference results are presented as follows: (Treatment Baseline / Treatment Midline, Control Baseline / Control Midline, Difference-in-Difference indicator). The DID indicator is followed by an asterisk if p-value is significant at the 10% level and by two asterisks if significant at 5%. Tables showing full difference-in-difference results across all analytical variables from the EM HHS, SVS and SBA as well as from projects' household surveys are shown in Annex G.

Box 5: Learning metrics and benchmarks used in EM data analysis

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

The **version of EGRA** 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 before 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 **version of EGMA** 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 (advanced addition, advanced subtraction, multiplication and division).

The EM aggregates learning data across the SCW using a single learning metrics for literacy and a single learning metrics 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. As for 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 before 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. Besides, 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.

Our unified approach enables us to assess the effects of projects on learning across diverse national and local contexts, ages, grades, abilities, types of interventions. But 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 than another, which may reduce the observable effect of projects aiming 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**. 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[50].

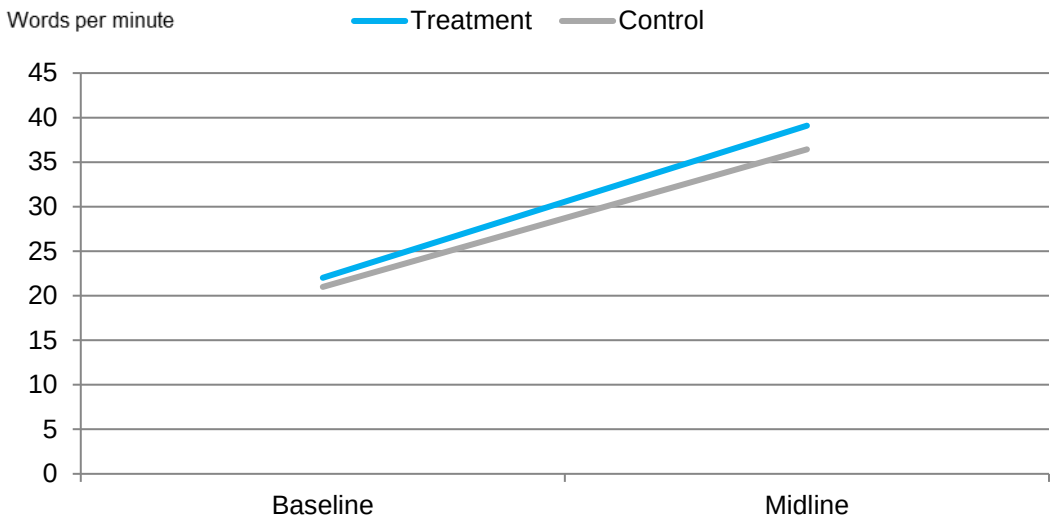
We refer to fluency benchmarks published by Abdazi (2001)[51] for use by the World Bank. Having tested students in US schools, Abdazi suggests to use their median score as a benchmark for students in high-income countries. Scores of second-graders at the 18th percentile (i.e. the lower end of the distribution) were similar to those suggested by USAID for use in developing countries, and we are therefore using the scores achieved by the 18th percentile in US schools across different grades as the benchmarks for GEC students of different ages.

To date, no comparable benchmarks have been developed for the assessment of EGMA results. There is no established, aggregate EGMA score that readily represents mathematical ability as accurately as oral reading fluency (in wpm) represents literacy across subtasks.

3.3.1 What impact has the SCW had on literacy?

At baseline, girls in the treatment group read on average 22 words per minute. At midline, average reading fluency scores increased to 39 words per minute. But reading fluency improved just as strongly in the control group and there is no significant difference in the trends observed for the two groups (22 / 39, 21 / 36, +1.6, see Figure 17). The graph does not include EM data from Camfed project areas where girls' learning outcomes were assessed in exactly the same way as in other projects' areas, with the exception that there were no control areas. The specific sampling design for Camfed areas is described in Annex E.

Figure 17: EGRA mean oral reading score across SCW EM cohort excluding Camfed areas



There is a small but significant change in *median* reading fluency scores. At baseline, the median was zero across SCW treatment areas, indicating that more than half the girls read zero words per minute at baseline. At midline, the median shifted to 27 wpm, implying that half the girls could read at least 27 words per minute. In the control cohort, the median moved from 0 wpm to 23 wpm which is a significantly lower increase than in the treatment cohort (0 / 27, 0 / 23, +4.0**). See Box 6 below for how to interpret median and mean differences.

Figure 1 in Annex H illustrates the shifts in the baseline and midline distributions of EGRA oral reading scores across treatment and control areas, as well as the shifts in their respective means and medians. The same annex

[50] 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.
[51] Ibid.

presents tables showing difference-in-difference indicators for most variables across the EM household and school-visit survey, as well as school-based assessment data, at window and project level.

Box 6: Difference between mean and median scores

When analysing a continuous quantitative variable such as the EGRA word-per-minute score, two statistics can be used: the mean and the median.

The **mean** corresponds to the average value of all observations. It is the most commonly used statistic as well as the easiest to interpret, but, it is very sensitive to large and extreme values. In EM data, the oral reading score has a skewed distribution that goes from 0 wpm to 200 wpm (a few girls scored higher than 200 wpm when 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 would be the same when 100 girls go from 0 to 10 wpm as when 10 girls go from 100 to 200 wpm.

Using the **median** provides a way of giving the same weight to the lowest and the highest scores (hence to lower and higher performing students). The median is the value that splits the sample in two equal parts: exactly half students scored lower than the median, and exactly half students scored higher than the median. If between baseline and midline the scores of the best-performing half have increased but the scores of the other half have remained the same, the median will not change (as opposed to the mean which is likely to increase a lot). This does not mean that the median gives a summary of the scores of the lowest-performing students. Instead, it gives information on the *median student* and in the case of the oral reading score can be interpreted as follows: "Half of the students can read at least [the median] words-per-minute." Hence median difference gives a better sense of population shifts while the mean could in theory result from a change in only one individual.

Note that calculating a median only makes sense in the case of a continuous indicator. For binary variables that can be expressed as proportions or percentages, the median can either be equal to 0 or to 1. If the mean is below 50 per cent, median is equal to 0. If the mean is above 50 per cent, median is equal to 1. If the mean is exactly equal to 50 per cent, median is equal to 0.5. In this case, the median always gives less information than the mean about the variable's distribution.

Findings from quantile regressions including the median score are described in details in Section 3.3.3.

Other EGRA Subtasks

At midline, girls in the SCW treatment cohort had improved their ability to sound letters and comprehend reading passages slightly more than girls in the control group. Girls in the treatment cohort correctly sounded 31 letters, on average, at midline compared with 19 at baseline, whereas girls in the control group improved from 19 to 29 letters (19 / 31, 19 / 29, +2.4*). Reading comprehension improved from a score of 1.0 to 1.6 questions correct which is significantly more than in control areas (1.0 / 1.6, 0.9 / 1.4, +0.1*). In both cases, we measure a difference-in-difference that is significant at the 10 per cent level. In contrast, the EGRA invented word subtask does not show any significant difference in the trends of the two groups (12 / 21, 12 / 20, +1.2).

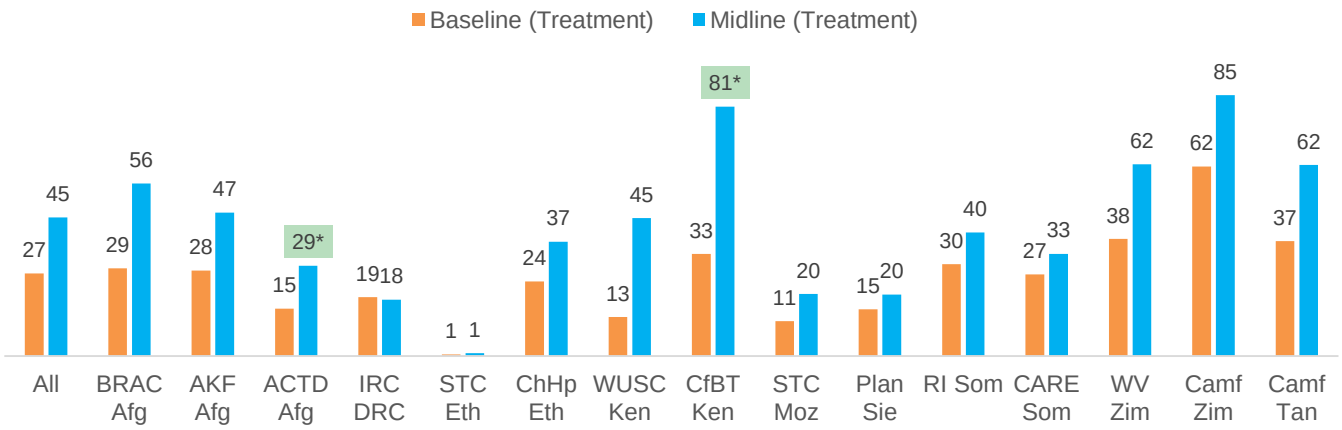
Oral reading fluency at project level

At project level we observe a significant improvement in average reading fluency scores in the treatment cohort compared with the control cohort in CfBT (Kenya) and ACTED (Afghanistan) project areas. Girls in CfBT project areas more than doubled their baseline scores, from 33 to 81 wpm (see Figure 18), improving their average score by 14 wpm more than girls in the control cohort (33 / 81, 35 / 69, +13.9**). An oral reading fluency score of 81 wpm indicates that the girls are now able to read unfamiliar texts fluently.

Girls in ACTED project areas improved their average scores from 15 to 29 wpm that is 5 wpm more than girls in the control group (15 / 29, 17 / 26, +4.8*). They also showed a significant improvement in reading comprehension, over and above that achieved in the control group (0.8 / 1.9, 1.0 / 1.7, +0.4*).

We find a significant difference-in-difference in *median* scores for BRAC (Afghanistan), WUSC (Kenya) and Plan (S-L) project areas, suggesting that they may have had a particular effect on girls towards the bottom of the score distribution. As shown in Figure 18 the WUSC treatment cohort also shows the second largest improvement in average reading fluency scores between baseline and midline (+32 wpm), after CfBT (+48 wpm), followed by BRAC (+25 wpm).

Figure 18: Oral reading fluency by project area (treatment cohort)



The data shows virtually no improvement in scores for girls in IRC (DRC) and Save the Children (Ethiopia) project areas. In Save the Children (Ethiopia) project areas, the average girl in the treatment cohort is still unable to read more than one word per minute at midline. Figure 18 also illustrates the stark variation in reading fluency across the window: Pastoralist girls in Save the Children (Ethiopia) project areas are still largely illiterate, whereas girls in Camfed project areas in Zimbabwe (who are older on average) already had a basic level of fluency at baseline, and were relatively fast readers at midline.

Oral reading fluency by subgroup

Figure 19 breaks down the change in reading fluency scores by school phase and age group for girls in the treatment cohort. Out-of-school girls show a far steeper improvement in oral reading fluency than any other group (29 words per minute). It is the only group for which we measure a significant difference-in-difference when comparing girls in the treatment and control group (9 / 38, 3 / 20, +12**). This suggests that the GEC has been particularly effective at improving literacy skills for out-of-school girls, at least compared to other interventions taking place in control areas. Among girls enrolled in upper primary school, in contrast, we find almost no change in reading fluency between baseline and midline.

Figure 19: Oral reading fluency by school phase and age group (treatment cohort)

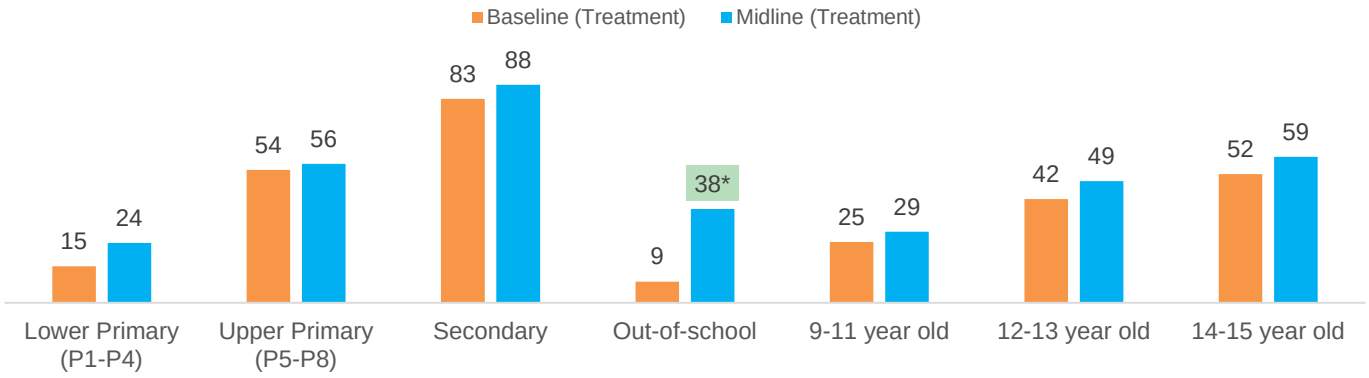
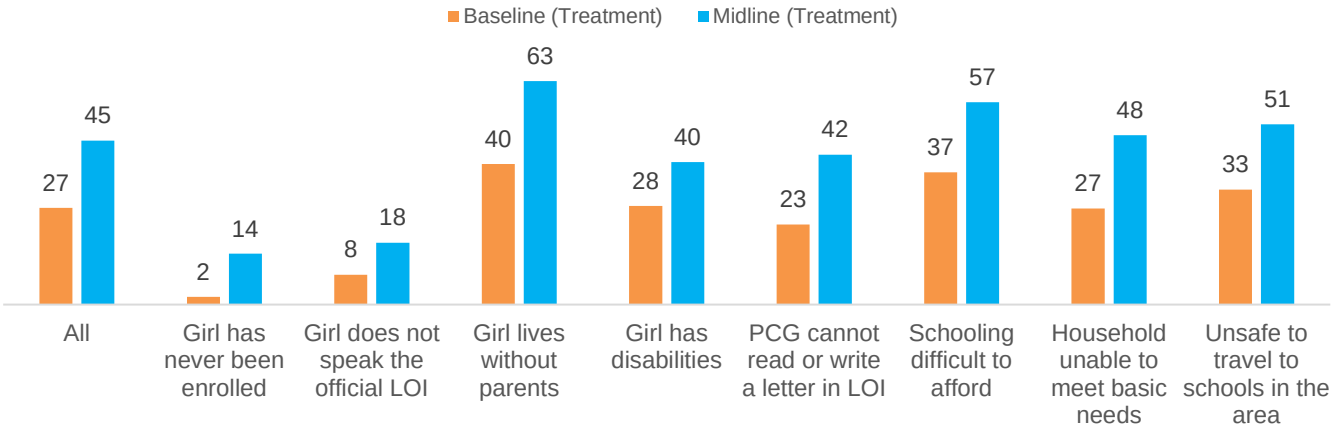


Figure 20 shows changes in oral reading fluency for specific groups of marginalised girls in the treatment cohort. The average improvement across the cohort is an increase of 19 wpm. For most subgroups, the increase is roughly similar, although they started from different levels at baseline. We observe the smallest increase (by 10 words per minute) among girls who do not speak the language of instruction. They read only slightly faster than girls who have never attended school. This highlights the particular barriers faced by girls who are not fluent in the language of instruction and suggests that their disadvantage has persisted since baseline.

Figure 20: Oral reading fluency by subgroup at baseline and midline (treatment cohort)



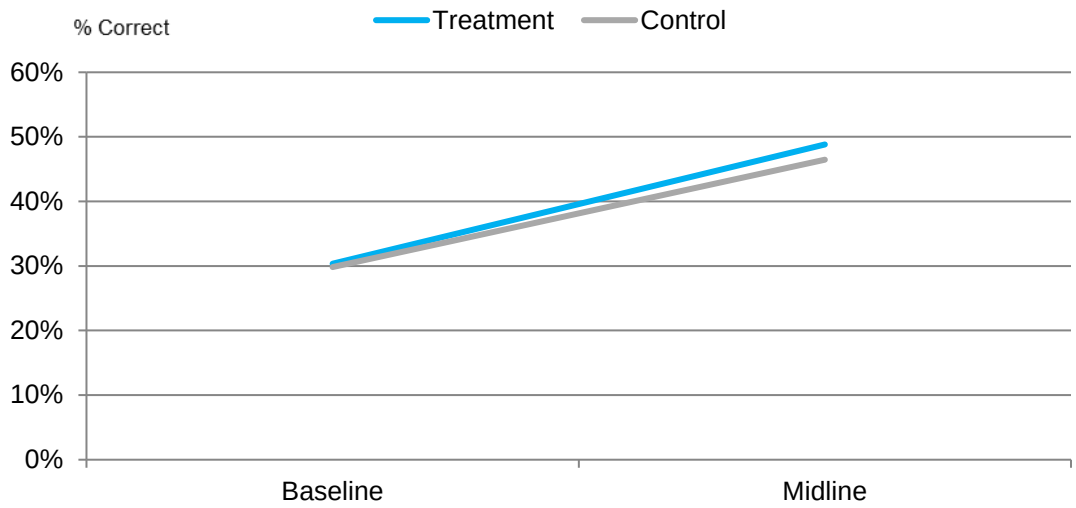
However, when looking at the median of the distribution, we find a significant effect on a number of subgroups. We measure the largest difference-in-difference for out-of-school girls, whose median reading fluency has improved by 19 words per minutes over and above the median change in the control group. We measure an almost equally large shift for girls with low attendance levels (19 wpm), girls whose families struggle to afford the cost of schooling (12 wpm), and girls whose parents cannot read or write a letter in the language of instruction (11 wpm).

This supports the hypothesis that SCW interventions were particularly effective in improving reading fluency of low performing students. This is only partly confirmed by our analysis of the different EGRA subtasks. As described earlier, GEC activities seem to have had a positive net effect on the scores of letter sound and reading comprehension subtask, which are supposed to be the easiest and the most difficult subtask of EGRA respectively (although those effects are only significant at the 10 per cent level and hence quite small). But the two middle subtasks, invented word and oral reading, do not show any significant improvement over-and-above the control group.

3.3.2 What impact has the SCW had on numeracy?

At baseline across the SCW, girls in the treatment cohort scored on average 30 per cent correct answers on the EGMA test. At midline, their average scores had improved to an average of 49 per cent. But as illustrated in Figure 21, the improvement was almost exactly the same in the treatment group as in the control group and we do not measure any significant GEC effect on numeracy at the SCW level (30% / 49%, 30% / 46%, +2 %-points).

Figure 21: EGMA percentage correct of all sub-tasks, SCW EM cohort excluding Camfed areas

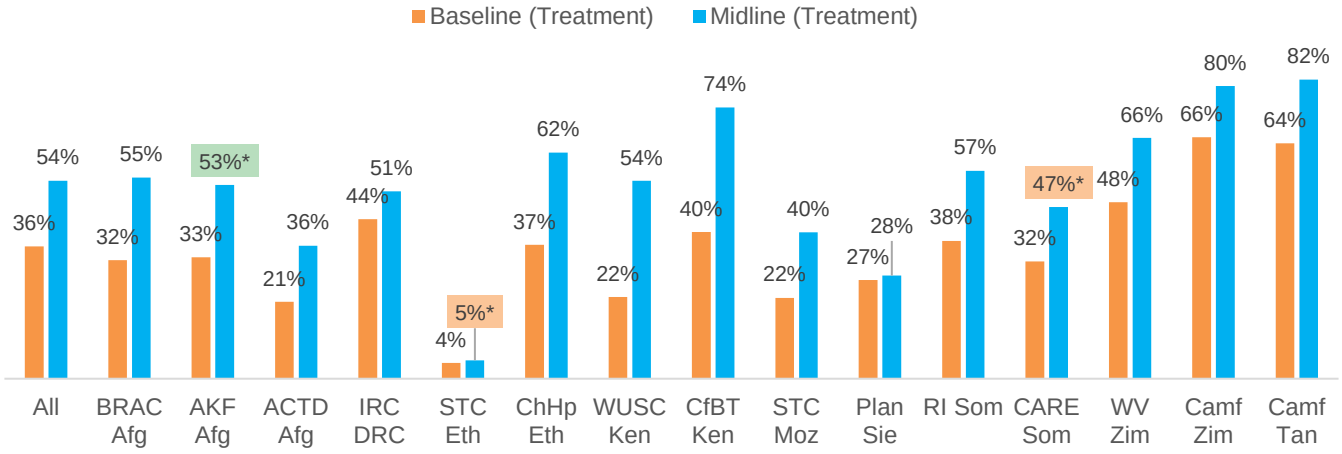


At project level using our EM data, we measure a significant positive effect on EGMA scores in Aga Khan Foundation (Afghanistan) project areas. Girls in the treatment cohort improved their average score from 33 per cent to 53 per cent, which is 8 percentage points more than the improvement measured in the control cohort (33% /

53%, 29% / 41%, +8* %-points). As shown in Figure 22 we measure almost no improvement in numeracy in Save the Children (Ethiopia) project areas, where girls scored only 4 per cent correct at baseline and midline. In fact, the control cohort improved their numeracy scores significantly more than the treatment sample (4% / 5%, 8% / 14%, -5** %-points). We also measure a negative difference-in-difference in CARE (Somalia) project areas. While there was some improvement between baseline and midline in the treatment cohort, the control improved their average score significantly more (32% / 47%, 28% / 52%, -9* %-points).

With the exception of Save the Children (Ethiopia) and Plan (Sierra-Leone) project areas, numeracy scores improved substantially across the window, closely following the patterns observed for oral reading fluency. We measure particularly large improvements from baseline to midline in CfBT (Kenya) and WUSC (Kenya) project areas but there is no significant difference-in-difference with the control group. As in the case of reading fluency, Camfed (Tanzania and Zimbabwe) project areas show the highest absolute EGMA scores.

Figure 22: Numeracy (EGMA) scores by project area (treatment cohort)



Numeracy by subgroup

When looking at changes in EGMA scores by school phase and age (see Figure 23), there is an increase for girls in lower primary school, and a slight decrease in average scores among girls in upper primary and secondary school. We find a particularly large improvement in numeracy among out-of-school girls, who improved their EGMA scores from an average 15 per cent at baseline to 45 per cent at midline. Similarly to reading fluency, out-of-school girls are the only subgroup for which we find a significant difference-in-difference compared with the control group (12% / 46%, 7% / 27%, +13** %-points). This suggests that the GEC has had a significant effect on the numeracy levels of out-of-school girls in treatment areas.

Figure 23: Numeracy (EGMA) scores by school phase and age group (treatment cohort)

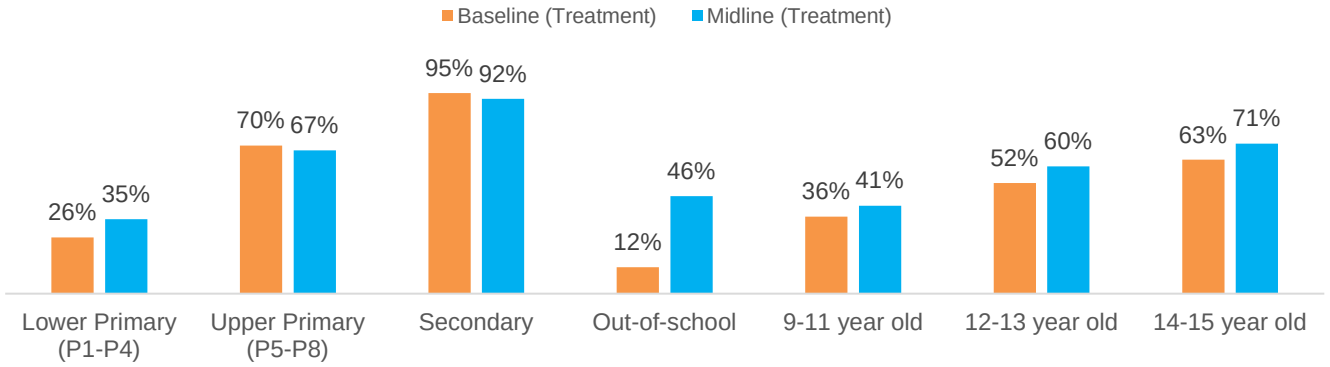
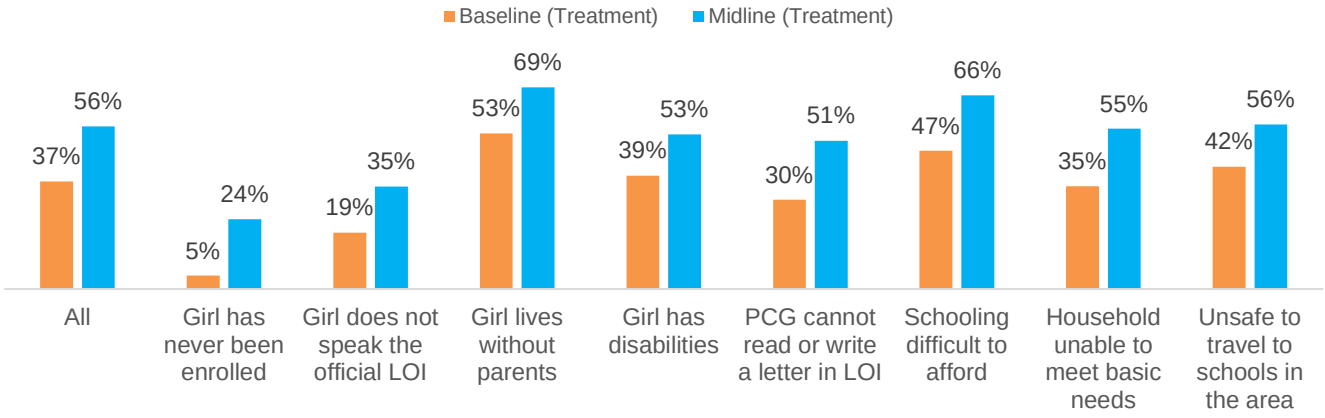


Figure 24 looks at EGMA scores at baseline and midline for specific groups of girls in the treatment cohort. As could be expected, girls who have never attended school scored lowest at baseline, followed by girls who do not speak the official language of instruction, and girls whose primary caregiver cannot read or write a letter in the

language of instruction. For all three groups, we see improvements by roughly 15 percentage points from baseline to midline. Girls living without their parents scored proportionately higher, but this is linked to the fact that many of these girls are from Camfed (Tanzania and Zimbabwe) or World Vision (Zimbabwe) project areas where the samples had a higher number of older girls, with higher average numeracy scores. None of these subgroups show a significant difference-in-difference with the control group, except for girls living without their parents, who whom we measure a significantly smaller increase in EGMA scores compared with the control group.

Figure 24: EGMA scores at baseline and midline by population subgroup (treatment cohort)



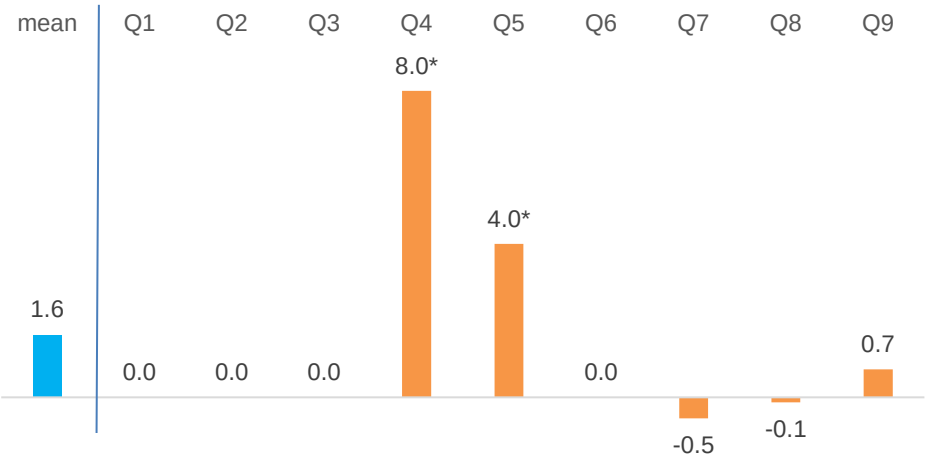
3.3.3 Difference-in-difference on deciles of the learning scores' distribution

EGRA (oral reading score)

Figure 25 below shows the difference-in-difference (DID) indicators for the EGRA oral reading score, expressed in words-per-minute, between baseline and midline across all SCW project areas excluding Camfed project areas. DID regressions were run separately based on the mean of the oral reading score distribution as well as on the nine deciles.

Deciles split the data into ten equal parts. The higher a girls' overall reading fluency score, the higher she will be placed in the decile ranking. For example, if the girl scored higher than 90% of all the tested girls then she would be in the 9th decile (Q9). A girl whose score was amongst the lowest 10% would be placed in the first decile (Q1). The 5th decile (Q5) is also called the **median** (see Box 6), that is the score value at which the distribution is split within two equal parts. To give another example, if the 70th percentile of the oral reading score distribution was 60 words per minute, then a girl scoring 60 words would score better than 70% of the test-takers.

Figure 25: DID indicators for oral reading score across all SCW projects' areas (excluding Camfed)



As mentioned above, the difference between the mean change in reading fluency scores measured in the treatment and control groups between baseline and midline (i.e. the difference-in-difference) was equal to +1.6

words per minute across the SCW (22 / 39, 21 / 36, +1.6). This means that girls in the treatment group improved their scores, on average, by 1.6 words more than girls in the control group.

Figure 25 shows that the impact of the GEC has been concentrated around the median of the distribution: among girls placed in the 4th decile (i.e. scoring better than 40% of girls, but not better than 50%), reading improved by 8 words per minute more than in the control group between baseline and midline (0 / 14, 0 / 6, +8.0**). The median score (i.e. the mid-point of the distribution at which half the girls scored lower, and half scored higher) augmented by 4 words per minute more in the treatment than in the control group (0 / 27, 0 / 23, +4.0**). Other deciles show non-significant difference-in-difference indicators, all of them lower than one word-per-minute. These figures suggest that **the SCW impacted most literacy around the middle of the distribution, rather than on the lowest or highest performing girls in the cohort.** These window-level results are confirmed by further analysis at the project level, as shown in Table 12.

Table 12: Difference-in-difference indicators of the oral reading score (in wpm), regressed on means and on deciles

	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
mean	1.6	0.3	4.8	4.8	-0.3	-0.9	0.7	0.6	13.9	-1.8	0.0	6.1	-7.0	-0.5
Q1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	-1.0	0.0	0.0	0.0	0.0	0.0
Q2	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	24.0	0.0	0.0	0.0	0.0	-13.0
Q3	0.0	10.0	19.0	5.0	0.0	0.0	-4.0	3.0	25.0	0.0	0.0	15.0	-13.3	11.0
Q4	8.0	11.0	13.0	2.0	0.0	0.0	-4.0	9.0	12.1	0.0	0.0	19.6	-20.9	1.0
Q5	4.0	15.0	-1.0	5.0	0.0	0.0	11.0	8.0	8.6	0.0	5.0	-6.0	-24.0	-8.0
Q6	0.0	2.0	0.0	7.0	5.0	0.0	5.5	2.0	19.6	-2.0	4.0	-9.0	-15.0	1.7
Q7	-0.5	-15.0	11.0	3.0	4.0	0.0	1.0	-2.0	28.7	1.0	-3.0	-4.0	-4.0	-1.0
Q8	-0.1	-8.0	11.0	4.0	-29.8	0.0	7.0	9.4	22.3	-0.8	-2.0	2.0	8.0	3.2
Q9	0.7	5.8	2.3	10.0	7.6	-5.0	-1.9	-16.2	5.2	-8.0	9.8	9.9	-7.3	-8.7

Note: Q5, the 5th decile, is equivalent to the median. Regressions are based on the cohort sample only, which explains why some large difference-in-difference (such as the 8th decile in IRC) are not statistically significant as they are based on smaller sample sizes.
Green (resp. orange) shows statistical significance at the 10% level of a positive (resp. negative) difference-in-difference indicator.

Most projects achieved significant difference-in-difference indicators around the third, fourth and fifth deciles (the median), which suggests than the lower-median children have been those benefitting the most from the GEC SCW interventions. Only in CfBT (Kenya) areas our EM data shows a positive effect in deciles Q2, Q7 and Q8, and only in BRAC (Afghanistan) areas we observe a significant positive effect in the top decile Q9. We do not observe any positive difference-in-difference indicator for the first decile Q1 (i.e. the lowest scoring girls).

It is important to note that these results are largely influenced by the metrics used. In particular, floor and ceiling effects may prevent the detection of change in scores for the more extreme categories corresponding to the lowest and highest deciles (we capped the maximum possible score at 200 words, after accounting for time remaining upon completion of the exercise).

In all SCW project areas except for CfBT (Kenya), the lowest decile Q1 was equal to zero at baseline and at midline. This means that at least one tenth of girls scored zero words-per-minute at both waves, and explains why we measure no positive effect for this lowest decile. The same applies, although to a slightly lesser extent, to the second-lowest decile Q2, and to most of the score distribution in Save the Children (Ethiopia) project areas where girls up to the eight decile scored zero words-per-minute (which implies that less than 20% of cohort girls could read more than a single word at baseline and midline).

Similarly, ceiling effects make it difficult to capture the impact of the SCW on the highest performing girls, although this should only be a minor limitation as only a few girls reached the maximum score of 200 words-per-minute accounting for time remaining[52].

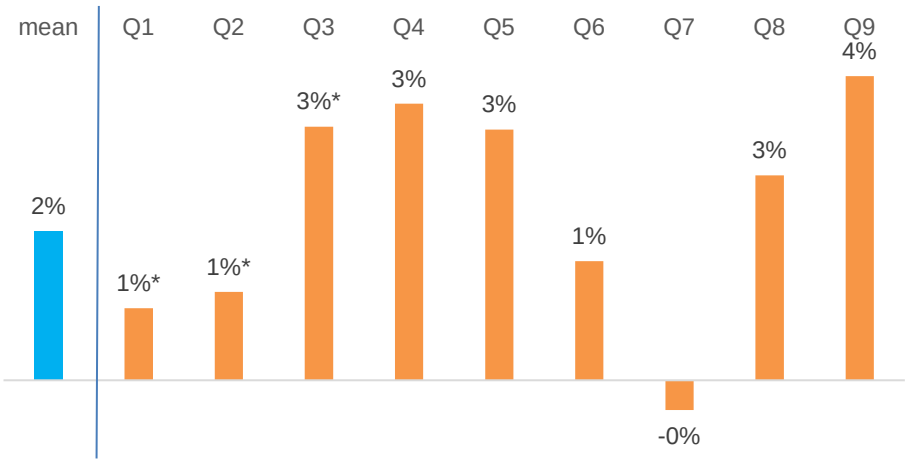
[52] Note that the number of items in the EGRA oral reading score is usually around 60 or 65. This implies that the difference of scores between students who scored more than 60 are based on time remaining only. Among highest performers, a positive difference-in-difference therefore captures a greater rapidity rather than a greater reading fluency.

EGMA (% correct across all subtasks)

We ran the same calculations on the EGMA score, defined as the percentage of items solved correctly across all EGMA subtasks. The figure below shows the DID indicators for the mean and the nine deciles, expressed in percentage points. Deciles Q3, Q4 and Q5 (the median) show the highest effects in absolute terms, but only the effect on decile Q3 is statistically significant[53].

Figure 26 shows that **the overall effect of SCW projects on EGMA scores has tended to focus on the lowest-performing girls**, rather than on the medium deciles as we observed on the EGRA oral reading score.

Figure 26: DID indicators for % EGMA correct across all SCW projects' areas (excluding Camfed)



This is confirmed by the project-level results shown in Table 13 below, where the effect is concentrated on the first three deciles, plus the fourth decile for WUSC (Kenya). Only in ACTED (Afghanistan) project areas we see a larger effect on the medium deciles Q5 and Q6. AKF (Afghanistan) is the only project area where we find a positive and significant effect on the mean as well as on the highest-scoring girls in decile Q9.

Table 13: Difference-in-difference indicators of the % EGMA correct, regressed on means and on deciles

	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
mean	2%	1%	8%	2%	4%	-5%	2%	5%	4%	0%	4%	4%	-9%	-1%
Q1	1%	2%	5%	0%	-1%	0%	0%	2%	1%	-2%	-1%	4%	-1%	1%
Q2	1%	9%	8%	-2%	-8%	0%	6%	8%	3%	-1%	-1%	12%	-9%	-1%
Q3	3%	11%	15%	-2%	5%	0%	8%	15%	-1%	2%	0%	14%	-16%	-7%
Q4	3%	4%	1%	5%	3%	0%	7%	14%	2%	1%	-2%	-6%	-7%	-5%
Q5	3%	-9%	0%	11%	2%	-1%	7%	13%	13%	-3%	6%	-13%	-12%	-4%
Q6	1%	-6%	9%	10%	3%	-1%	-3%	7%	7%	3%	10%	-13%	-4%	0%
Q7	0%	-11%	10%	2%	6%	-6%	-5%	-1%	9%	1%	4%	8%	-13%	4%
Q8	3%	-1%	6%	0%	18%	-15%	2%	-1%	-4%	0%	12%	11%	-10%	8%
Q9	4%	4%	19%	0%	4%	-27%	2%	-1%	3%	2%	9%	6%	-7%	-1%

3.3.4 Outcomes by exposure

Further analysis (see Table 14) explores changes in learning outcomes based on whether or not girls have been exposed to specific GEC interventions. We use variables from the household survey and from the school visit survey to assess whether girls have been part of any GEC-type interventions.

[53] The lowest deciles Q1 and Q2, although demonstrating a smaller net effect as Q4 and Q5, also show statistical significance. This is easily explained by the fact that scores in the lowest quintiles are much smaller than scores in the medium and highest quintiles. Hence a small effect (in absolute terms) can be statistically significant while a large effect in higher quintiles can be non-significant.

- Girls attending special classes or study groups improved their oral reading fluency by 20 wpm, compared with 17 wpm among girls who did not participate in such activities. The difference of 3 wpm is statistically significant at the 10 per cent level.
- Girls who received special tutoring or help with their school work improved their reading fluency scores by 25 wpm, compared with 16 wpm among girls who had not received such support, according to data from the school visit survey. The difference of 9 wpm was statistically significant at the 5 per cent level.

The school visit survey also allows us to look at whether or not learning outcomes have improved depending on activities that took place in school. It should be noted, however, that this data does not cover Afghan, Plan and Camfed project areas were school-based data collection was not possible at midline.

- Girls attending schools where activities to improve learning have taken place since baseline have improved their reading fluency score by 19 wpm, compared with 13 wpm among girls whose schools did not carry out such activities. The difference of 6 wpm was statistically significant.
- Where teachers had received training in either general teaching methods, or in gender-sensitive pedagogy, girls improved their reading by 7 wpm more than girls at schools where no such training had taken place. The difference was statistically significant for both types of training.
- Where teachers had received training in gender-sensitive pedagogy, girls improved their oral reading fluency by 19 wpm, compared with 15 wpm among girls enrolled in schools were no such training had taken place. The difference of 4 wpm was smaller than that measured for other types of teacher training, but still statistically significant.

Table 14: Change in EGRA from baseline to midline depending on reported exposure in treatment areas

			Received school books	Had a bursary		Attended special classes / study groups		Received special tutoring / help with homework		Activities to improve learning at school	Teachers were trained in:			
											General teaching methods	Gender sensitive pedagogy	Special needs education	Mentoring or coaching
			HHS	HHS	SVS	HHS	SVS	HHS	SVS	SVS	SVS	SVS	SVS	SVS
Treatment	No	Mean	18	18	18	17	17	18	16	13	12	15	15	17
		N	1,944	1,865	831	1,931	790	1,968	758	259	237	384	658	479
	Yes	Mean	17	18	12	20	20	19	25	19	19	19	22	18
		N	221	300	103	234	144	197	176	670	686	526	262	431
	Difference	Mean	-1	1	-6	3	3	1	9	5	7	4	7	0

Note: Light green shows statistical significance at the 10% level. Darker green shows statistical significance at the 5% level.

We find similar patterns for EGMA scores as for EGRA reading fluency scores. Girls who were exposed to tutoring or homework support, or teacher training all improved their EGMA scores significantly more than girls who had not been exposed to such interventions. We find similar patterns in the control group, which suggests that being exposed to these types of interventions supports learning regardless of whether or not they are GEC-specific interventions.

We generally find a **closer link between exposure to interventions and improved outcomes when using evidence from the school visit survey**, than when using responses from the household survey. This may be due to the fact that teachers and school administrators reported more factually on activities happening at their school, whereas caregivers may have reported more subjectively based on their perceptions and information passed on through word of mouth.

We do not measure any significant difference for several other exposure questions, such as whether or not somebody has spoken to the girl about enrolling or staying in school; whether any NGO or religious organisation has supplied funds to the school in the two years since baseline; and whether the school conducts any activities specifically to support girls. We also found no difference between girls who have been provided with textbooks or bursaries, and girls who have not.

This analysis logically suggests that interventions that aim at improving learning directly, such as tutoring, learning support or teacher training, have had a more direct and quick impact on girls' learning than indirect "activities" such as the provision of school materials and scholarships. These interventions may help girls go to and stay in school

as well as create better conditions for learning, but they are not sufficient to ensure girls are effectively improving their literacy and numeracy skills.

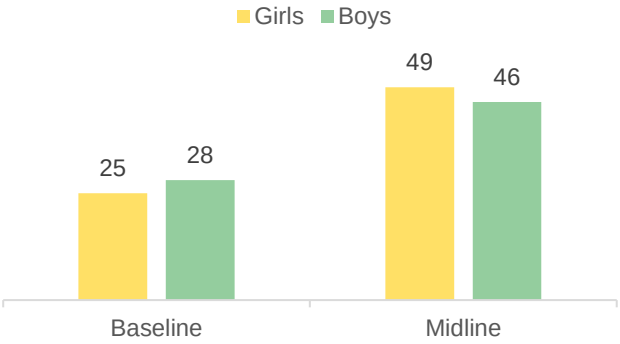
3.3.5 Gender disparity in learning outcomes

Literacy

To assess gender gaps in learning, the EM tested boys and girls in randomly selected treatment and control schools in Ethiopia, Kenya and DRC (in five project areas). At baseline, these children were in the second and fourth grade. Two years on, at midline, the same children were tested to the extent possible, and if not were replaced with children of the same gender and approximate grade level (i.e. fourth and sixth grade)[54].

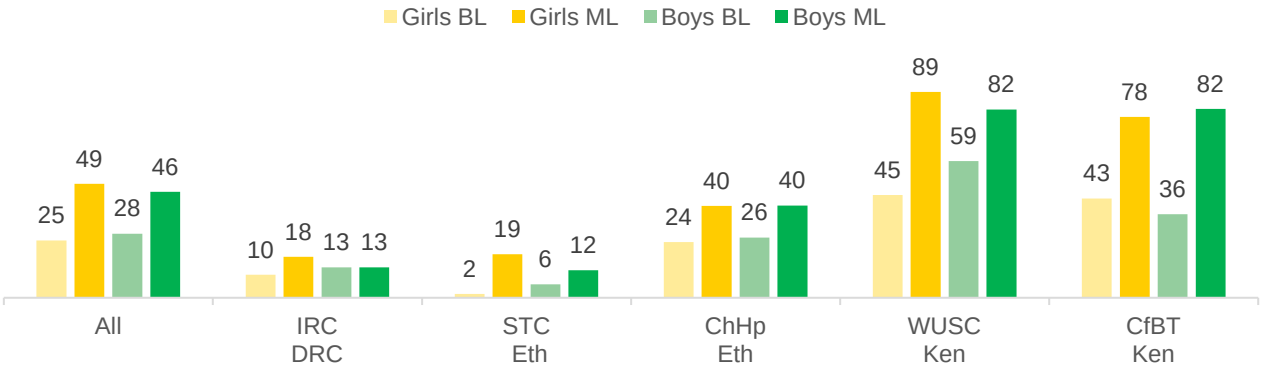
At baseline, girls in the treatment group read slightly fewer words per minute than boys (25 wpm compared with 28 wpm, see Figure 27). At midline, the trend has reversed with girls reading a few more words per minute (49 wpm) than boys (46 wpm). Girls' literacy scores have improved more (+25 wpm) than boys' scores have (+18 wpm) in the treatment group, suggesting that girls are actually learning faster than boys. This could in theory be explained by the effect of GEC interventions targeting girls specifically, but a similar trend is observed in control areas (girls improved by +26wpm and boys by only +20 wpm).

Figure 27: Oral reading fluency for boys and girls in grades 4 and 6 (treatment group only)



As shown in Figure 28, we observe the above-mentioned pattern of girls overtaking boys in three project areas. Girls are at a par with boys at midline in ChildHope (Ethiopia) project areas, and remain slightly behind boys in CfBT (Kenya) project areas. In most project areas, gender gaps are small (around 2 to 5 wpm). This aligns with our finding at baseline that girls are not systematically more educationally marginalised than boys in all SCW target areas. Indeed, the data suggests that boys might be at a disadvantage compared with girls in a number of SCW contexts, at least in grades 4 and 6, which were tested.

Figure 28: Oral reading fluency for boys and girls (treatment group only), by project



As in the household survey, we see the largest improvements for girls and boys in WUSC (Kenya) and CfBT (Kenya) project areas (see Figure 28). Girls in IRC (DRC) project areas improved their reading fluency slightly,

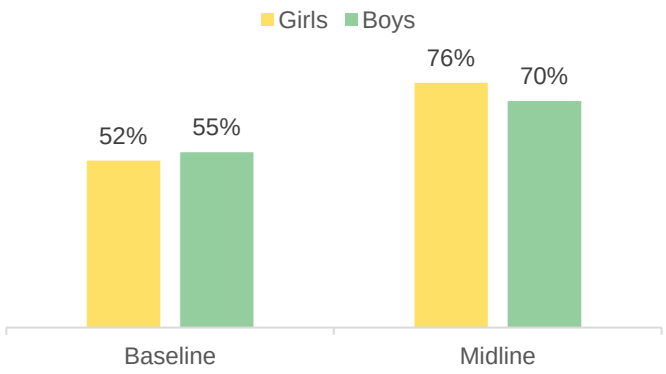
[54] In the SBA, children were not given identification numbers at baseline. Therefore, it is not possible to determine the level of attrition between baseline and midline and to determine whether there have been any systematic, gendered differences in the types of children who dropped out of the sample. However, we did not observe any systematic gender gap in learning outcomes at baseline, so it is unlikely that a gendered attrition bias towards lower-performing students would have biased our current analysis of the gender gaps between boys and girls.

whereas boys did not improve their scores from baseline. Interestingly, this pattern resembles that observed among girls in the full IRC (DRC) household survey sample. We measure no significant difference-in-difference for either boys or girls in any project area.

Numeracy

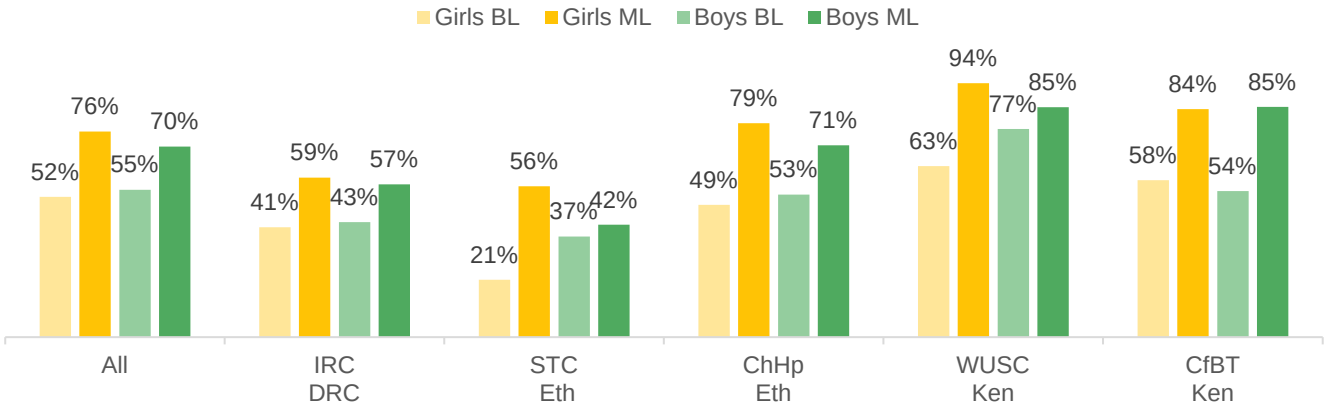
As shown in Figure 29, we observe a similar pattern for numeracy as for literacy, with girls starting off slightly behind boys at baseline and overtaking them by midline. Girls' EGMA scores improved by 24 percentage-points, whereas boys' scores improved by only 16 percentage-points. Again, we find no significant difference-in-difference when comparing changes in the treatment group with changes in the control group.

Figure 29: Numeracy (EGMA) scores for boys and girls in grades 4 and 6 (treatment group only)



As shown in Figure 30, girls overtook boys in numeracy in all project areas, except CfBT where they started off slightly ahead of boys at baseline and reached roughly the same score at midline. We observe an equally large improvement in girls' EGMA scores in save the Children (Ethiopia) project areas as we do for reading fluency, whereas boys have improved relatively little in comparison. However, we measure no significant difference-in-difference for either boys or girls in any project area.

Figure 30: Numeracy (EGMA) scores for boys and girls (treatment group only), by project



Our analysis of gender gaps using data from the school based assessment of both boys and girls confirms our baseline findings that girls are not systematically educationally disadvantaged, compared with boys. Even though girls started off slightly behind boys at baseline (when they were in grades 2 and 4), they have since slightly overtaken boys in a majority of project areas.

3.3.6 SCW projects' learning achievements

Table 15 below shows the projects' impact in literacy using projects' own data as shown in the final outcome spreadsheets.

Table 15: SCW projects' midline achievements in literacy – based on projects' data reanalysed by the FM

Literacy	BRAC- C	BRAC-G	AKF	ChHpe	WUSC	WV	IRC	STC	STC	CfBT	RI	CARE	Camfd	Camfd	Acted	Plan
	Afg	Afg	Afg	Eth	Ken	Zim	DRC	Eth	Moz	Ken	Som	Som	Tan	Zim	Afg	Sie
Midline target in SD	0.3	0.3	0.3	0.3	0.2	0.2	0.3	0.3	0.25	0.25	0.2	0.2	0.3	0.3		
Data quality rating	Partly conclusive	Partly conclusive	Partly conclusive	Conclusive	Conclusive	Partly conclusive	Partly conclusive	Partly conclusive	Partly conclusive	Conclusive	Partly conclusive	Partly conclusive	Partly conclusive	Partly conclusive		
Learning test and unit	EGRA wpm	EGRA wpm	EGRA wpm	EGRA wpm	EGRA wpm	EGRA wpm	EGRA / 100	EGRA / 100	EGRA / 100	UWEZO	Adapted UWEZO	Adapted UWEZO	National	National		
Baseline (Treatment)	43.5	76.5	16.4	41.5	60.8	82.1	11.2	9.7	26.3	3.2	7.7	32.9	34.6	19.3		
Baseline (Control)	benchmark	benchmark	benchmark	40.0	benchmark	85.6	8.7	21.2	29.1	3.4	benchmark	benchmark	35.0	20.5		
Midline (Treatment)	60.3	89.1	32.6	52.1	93.7	149.1	36.1	17.1	54.7	3.9	9.7	63.7	50.3	34.5		
Midline (Cont./Bench.)	43.5	88.3	23.7	53.1	66.4	154.3	24.9	32.5	58.2	3.7	9.8	55.0	42.9	30.1		
Impact in % (unadjusted)	**218%**	**8%**	**240%**	**-27%**	**299%**	**-30%**	**160%**	**-74%**	**-9%**	**78%**	**-15%**	**178%**	**231%**	**115%**		
Impact in % (adjusted)	**218%**	**8%**	**240%**	**12%**	**299%**	**24%**	**162%**	**-153%**	**23%**	**102%**	**-15%**	**178%**	**215%**	**96%**		
Adjusted - unadjusted impact	0%	0%	0%	39%	0%	54%	2%	-79%	32%	24%	0%	0%	-16%	-19%		
Final judgement	Extraordinary impact	Marginal or no impact	Extraordinary impact	Marginal or no impact	Extraordinary impact	Marginal or no impact	High impact	Negative impact	Marginal or no impact	High impact	Negative impact	Very high impact	Very high impact	Moderate impact		
EM DID indicator (EGRA wpm score)	0.3	0.3	4.8	0.7	0.6	-0.5	-0.3	-0.9	-1.8	13.9	6.1	-7.0			4.8	0.0
Midline C / Baseline T	0%	15%	45%		9%						27%	67%				
(ML C - BL C) / BL T				32%		84%	145%	116%	111%	10%			23%	49%		

Description of the table

The first row of this table, "midline target in SD", shows the project's midline target in standard deviations, over and above the control group. This means that the difference-in-difference (DID) indicator must be equal to 0.2, 0.25 or 0.3 standard deviations for the project to show a 100 per cent impact. When the project does not have control groups, benchmarks are set on the basis of the scores obtained by higher grade students at baseline.

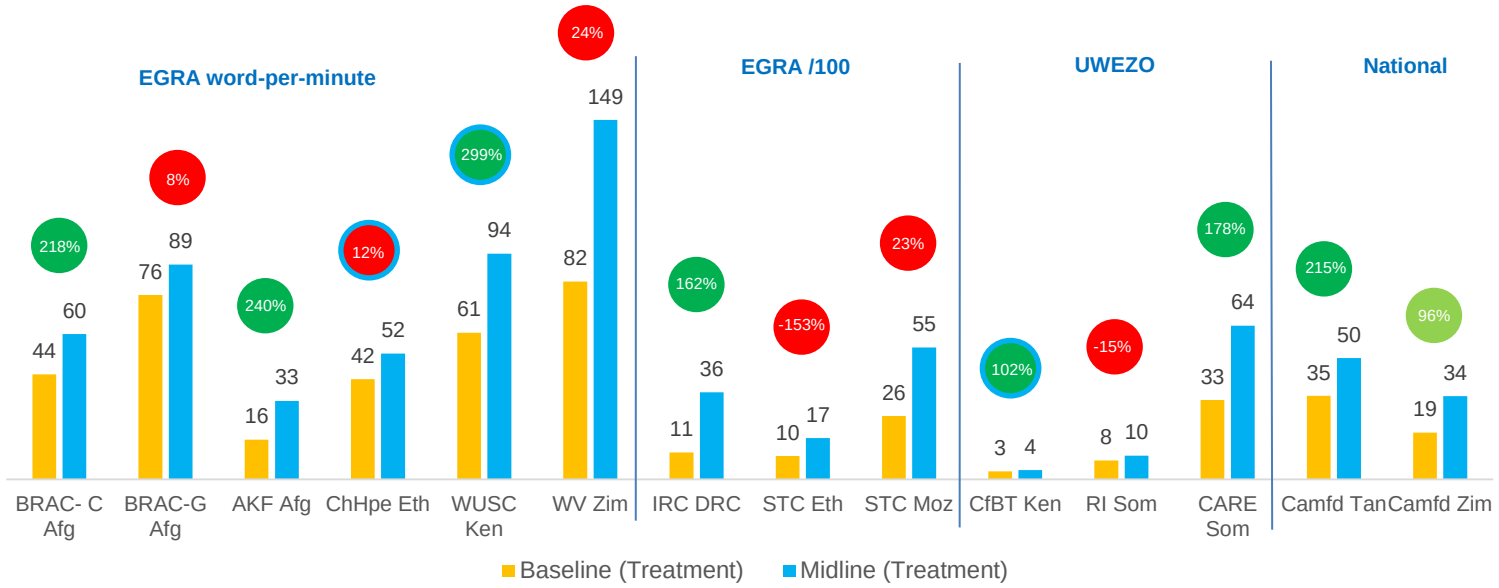
The second row, "data quality rating", is a general assessment made by the FM on the quality of the learning data submitted by the project. It is shared between the literacy and numeracy scores. ACTED (Afghanistan) and Plan (Sierra-Leone) did not submit any learning data at midline.

The third row shows the type of learning assessment and unit used. We have used this row to organise the columns. There are two columns for BRAC (Afghanistan), as community-based schools and government schools are assessed separately. Similarly, Camfed's project areas in Tanzania and Zimbabwe are considered independently from each other.

The following rows show baseline and midline scores for the treatment and control groups as weighted averages extracted from the final outcome spreadsheets validated by the FM. 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. For projects using control groups, this figure was obtained through a difference-in-difference regression analysis. The FM then adjusted this regression refining the model or using control and clustering variables, in order to assess for the specificities of each project's evaluation design. The difference between the adjusted and unadjusted impact figures is shown below, in green when strictly positive (the adjusted impact is larger than the unadjusted impact) and in red when strictly negative. The "final judgement" row below is a translation in qualitative terms of the adjusted impact in %. Figure 31 below shows aggregated baseline and midline literacy scores for the treatment group as well as projects' final impact in per cent.

Figure 31: SCW projects' final impact and aggregated literacy scores for the treatment group



The "EM DID indicator" shows the difference-in-difference indicator of the EM cohort sample based on the EGRA oral reading score. It is not supposed to be directly comparable to project's achievements. As opposed to projects' data, EM data is not supposed to be representative of each project's beneficiary population. Besides, the EM sample has not been designed to be disaggregated at the project level (see Box 3). We include it only as a sense-check of projects' achievements. EM DID indicators are shown in dark green when positive and significant at the 10 per cent level, in light green when greater than 5 words-per-minute but not significant, in red otherwise.

The last two rows come from EM calculations and have been included for indicative purposes only. They give indications on how much scores for the treatment group had to increase between baseline and midline in order to lead to a positive difference-in-difference (a larger difference in the treatment group than in the control group). The two rows hence give a proxy of how much change in scores was expected from the project. Greener cells shows projects for which a small change was expected, orange cells shows projects for which a large change was expected.

The first of the two rows shows the ratio of the midline control score over the baseline treatment score. It is relevant to projects who do not have control groups and for which benchmark scores were used. The second row shows the ratio of the difference between midline control and baseline control scores, over the baseline treatment score. It is relevant to projects using control groups. Note none of these two rows depend on the midline target set in standard deviations. An example of how to read it given below: “In order to achieve a positive impact, WUSC treatment scores had to increase by 9 per cent between baseline and midline.” All other things being equal, it was therefore much easier for WUSC (Kenya) to achieve a positive impact than for CARE (Somalia) for instance, which had to increase their scores by 67 per cent between baseline and midline in order to show a bigger change than in the control group.

Interpretation

According to the FM’s reanalysis of project data, seven projects out of 14 have achieved a high, very high or extraordinary impact in literacy. One has achieved a moderate impact. According to EM data, two projects have achieved a high impact and two projects have achieved a moderate impact. The only project to have achieved a high impact across the two sources of data is CfBT (Kenya). AKF (Afghanistan) and their consortium of partners have achieved an extraordinary impact according to their own data, and a moderate impact according to our data.

Among the eight projects measuring literacy using EGRA, only three have met their target. In contrast, across the five projects using tests other than EGRA, three have achieve a high impact and one, Camfed (Zimbabwe), has achieved a moderate impact (but very close to 100% impact).

Among the five projects using benchmarks, only one, CARE (Somalia), had a reasonably large difference to achieve, with the ratio of the midline control over the baseline treatment score being equal to 67 per cent. This number is still lower than the numbers of four projects using control groups (WV, IRC and the two STC’s projects). CARE still managed to achieve a high impact while BRAC (Afghanistan) and RI (Somalia) did not manage to despite “easier” targets.

Among the eight projects using control groups, four have had at least a high impact. It is worth noting that three out of these four projects had reasonably smaller increases to achieve, with the notable exception of IRC (DRC) which managed to achieve a high impact despite a tremendous increase of average scores in the control group, from 8.7 per cent to 24.9 per cent. These large changes could be due to characteristics of IRC’s target population, or to the specificities of the test and score used (although these tests should not differ too much from other projects using the same test and unit). This might also be due to poor sampling or incorrect administration of the test.

Notably, STC’s (Ethiopia) average scores differed substantially between treatment and control groups at baseline. This could be due to poor sampling, or to differences in the way tests have been administered in the two groups, which may have lowered the comparability between the treatment and control populations.

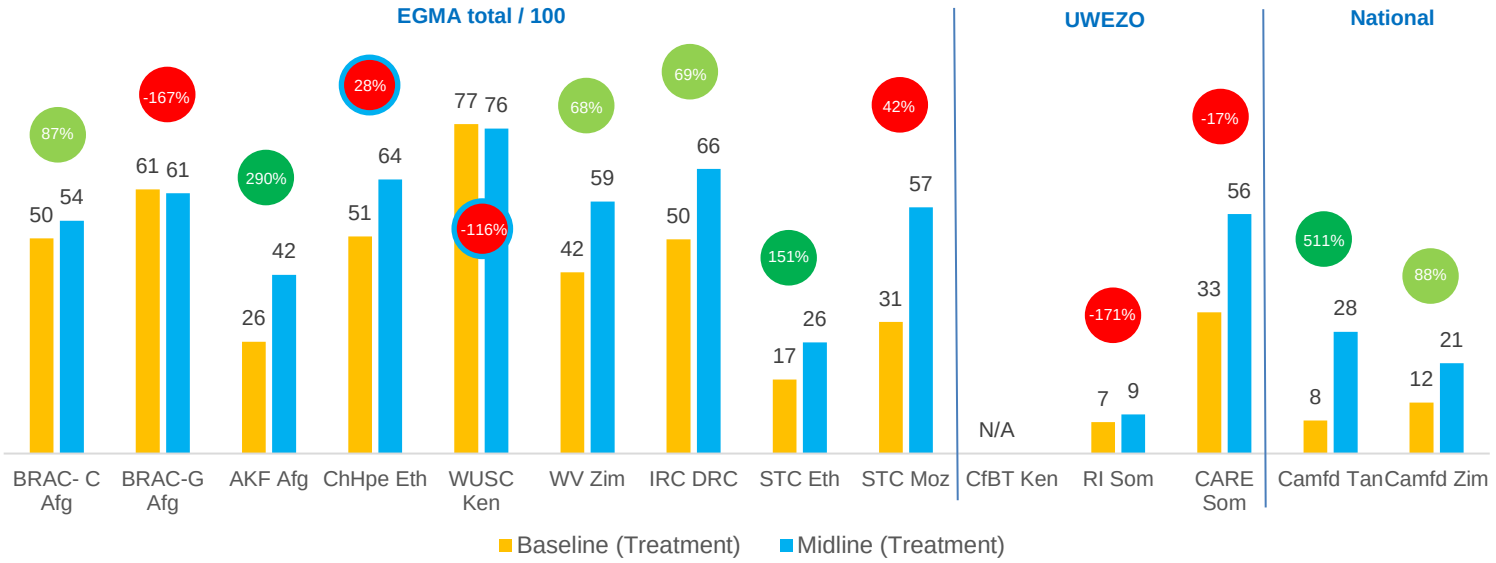
Table 16 below shows the projects' impact in numeracy using projects' own data from the final outcome spreadsheets.

Table 16: SCW projects' achievements in numeracy – based on projects' data reanalysed by the FM

Numeracy	BRAC- C	BRAC-G	AKF	ChHpe	WUSC	WV	IRC	STC	STC	CfBT	RI	CARE	Camfd	Camfd	Acted	Plan
	Afg	Afg	Afg	Eth	Ken	Zim	DRC	Eth	Moz	Ken	Som	Som	Tan	Zim	Afg	Sie
Midline target in SD	0.3	0.3	0.3	0.3	0.2	0.2	0.3	0.3	0.25	0.25	0.2	0.2	0.3	0.3		
Data quality rating	Partly conclusive	Partly conclusive	Partly conclusive	Conclusive	Conclusive	Partly conclusive	Partly conclusive	Partly conclusive	Partly conclusive	Conclusive	Partly conclusive	Partly conclusive	Partly conclusive	Partly conclusive		
Learning test and unit	EGMA / 100	EGMA / 100	EGMA / 100	EGMA / 100	EGMA / 100	EGMA / 100	EGMA / 100	EGMA / 100	EGMA / 100	UWEZO	Adapted UWEZO	Adapted UWEZO	National	National		
Baseline (Treatment)	50.0	61.4	26.0	50.5	76.6	42.2	49.8	17.2	30.6	N/A	7.3	32.9	7.7	11.8		
Baseline (Control)	benchmark	benchmark	benchmark	50.2	benchmark	42.4	46.6	43.2	31.7	N/A	benchmark	benchmark	9.4	13.0		
Midline (Treatment)	54.2	60.6	41.6	63.7	75.6	58.6	66.2	25.9	57.2	N/A	9.1	55.7	28.3	21.0		
Midline (Cont./Bench.)	50.0	66.3	34.3	62.5	79.1	58.2	59.5	37.8	57.5	N/A	10.2	56.6	12.3	19.9		
Impact in % (unadjusted)	**87%**	**-167%**	**290%**	**42%**	**-117%**	**49%**	**65%**	**197%**	**19%**	**N/A**	**-171%**	**-17%**	**543%**	**92%**		
Impact in % (adjusted)	**87%**	**-167%**	**290%**	**28%**	**-117%**	**68%**	**69%**	**151%**	**42%**	**N/A**	**-171%**	**-17%**	**511%**	**88%**		
Adjusted - unadjusted impact	0%	0%	0%	-14%	0%	19%	4%	-46%	23%		0%	0%	-32%	-4%		
Final judgement	Moderate impact	Negative impact	Extraordinary impact	Marginal or no impact	Negative impact	Moderate impact	Moderate impact	Very high impact	Marginal or no impact	N/A	Negative impact	Negative impact	Very high impact	Moderate impact		
EM DID indicator (EGMA % correct)	1%	1%	8%	2%	5%	-1%	4%	-5%	0%	4%	4%	-9%			2%	4%
Midline C / Baseline T	0%	8%	32%		3%						40%	72%				
(ML C - BL C) / BL T				24%		38%	26%	-32%	85%				38%	58%		

Figure 32 below shows aggregated baseline and midline numeracy scores for the treatment group as well as projects' final impact in %.

Figure 32: 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 project data re-analysis, and this is confirmed by EM data. According to the FM's project data reanalysis, three projects out of 13 have achieved a high, very high or extraordinary impact. Four of them have achieved a moderate impact. According to EM data, only one project, AKF and their consortium (Afghanistan), have achieved a significant impact in numeracy. It is also the project which has had the largest impact according to project data. CfBT's (Kenya) performance in numeracy has not been assessed quantitatively due to the incomparability of their test between baseline and midline (see Section 2.5.2).

Among the eight projects measuring numeracy using EGMA, only one has met its target: STC (Ethiopia). In literacy, STC's findings show a steep difference between scores from the treatment and control groups at baseline, which might be a hint that the two populations are not directly comparable. STC (Ethiopia) is also the only project to have a negative ratio of the difference between midline and baseline control scores over baseline treatment score (see last row). This is due to the fact that control scores have *decreased* between baseline and midline. As a result, even if treatment scores had not improved at all between baseline and midline, the project would have achieved a positive difference-in-difference.

The two projects in Somalia, which used adapted versions of UWEZO, both had a negative impact on numeracy. They also had larger increases to achieve between baseline and midline in order to meet their benchmarks compared to other projects that are using benchmarks. Camfed, which used a national test to assess girls' learning at baseline and midline, demonstrated an extraordinary impact of +511% in Tanzania and a moderate impact of +88% in Zimbabwe.

Among the five projects using benchmarks, four had a negative impact on numeracy and only BRAC (Afghanistan) managed to have a moderate impact. In literacy, in the absence of control group for this project, the benchmark used for the midline control score was the baseline treatment score. This implies that even the smallest increase in treatment scores between baseline and midline would have been sufficient to demonstrate a positive difference.

Among the seven projects using control groups, two have had at least a high impact in numeracy scores and three others have had a moderate impact. Among those five projects, two did not show any impact in literacy: World Vision (Zimbabwe) and Save the Children (Ethiopia). This may suggest that activities run by these two projects have had a bigger positive effect on girls' mathematics skills than on their reading skills.

Conclusion

The learning impact claimed by projects need to be assessed while considering the methodology used to calculate it. 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). However, our review suggests that targets before adjustment by the FM have been comparatively easier to reach for some projects than others.

This is especially the case for projects who dropped their control groups and for which benchmarks were used. Those benchmarks have been defined as the average baseline scores of the grades in which cohort girls were supposed to be enrolled at the midline stage. Considering the two-year interval between the baseline and midline data collection waves, those benchmark grades were often defined as the grades of the cohort at baseline plus two years. But in the case of BRAC community schools, the benchmark grade is the baseline grade. The measure of impact therefore relies on a simple before and after comparison. This potentially explains why BRAC community schools achieved an extraordinary impact of +218% in literacy and a moderate impact of 87% in numeracy (despite increasing from less than 10% in two years in the latter case).

The impact measured by projects who used control groups depends on the degree of comparability between the treatment and control groups at baseline. In the case of Save the Children (Ethiopia), it is striking that 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 evaluation design, and may partly explain why its literacy and numeracy achievements exhibit opposite patterns (-153% impact in literacy, +151% impact in numeracy).

It is worth noting that across the SCW, the simple difference between the treatment groups' scores between baseline and midline do not seem correlated with the final impact achieved by projects. Although this impact obviously depends on the target set in standard deviations, on the type of evaluation design used (control groups or benchmark scores) and on the adjustment calculated by the FM, one would expect a rough comparison between the baseline and midline treatment scores to give a hint about whether the project has achieved, or not, their target. We do not observe this. For example, World Vision (Zimbabwe) treatment literacy scores have almost doubled from baseline to midline (+82 wpm to +149 wpm) but the project has achieved only 24% of their target. Conversely, BRAC-C (Afghanistan), WUSC (Kenya) and Camfed (Tanzania) have all achieved an extraordinary impact higher than +200% but their treatment literacy scores only improved by about one third or less between baseline and midline. Similarly, Save the Children (Mozambique) treatment midline scores are about twice as big as at baseline for both literacy and numeracy. However, the project shows a marginal impact in both literacy and numeracy, achieving less than 50% of its targets.

There is little convergence between EM data and projects' data in terms of the measured impact. This is not surprising because they are based on different methodologies and, most importantly, the power of the EM sample is not large enough to draw any conclusions about the individual impacts claimed by each of the SCW projects. However, we assumed that a high or extraordinary impact would have been captured by our sample. We observe such a convergence for AKF (Afghanistan) and CfBT (Kenya). But EM data seems to contradict project's data in other areas: this is the case for IRC (DRC), BRAC-C (Afghanistan), WUSC (Kenya) and CARE (Somalia) for literacy, and Save the Children (Ethiopia) for numeracy. In particular, none of the high or extraordinary impacts claimed by projects who did not use control groups can be confirmed on the basis of EM data.

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

Enrolment and attendance rates have not significantly improved across SCW target communities compared to control areas. We showed at baseline that enrolment and attendance were already relatively high across the SCW, even though the proportion of out-of-school girls is much higher in some particular contexts (Somalia, Ethiopian Afar region, Kenya refugee camps) and for particular subgroups, especially secondary school age girls. Enrolment increased at midline in treatment areas but to the same extent as it increased in control areas. A large proportion of the observed increase is due to girls enrolling in school for the first time, especially among the youngest age groups, rather than to in-school girls being retained.

Attendance varies very little 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. Our school-visit survey now includes a refined set of variables to measure attendance, so we hope to be able to provide a more valid midline to endline comparison. The measure presented at midline suggests that midline attendance levels are relatively high across the SCW with girls attending 89 per cent of school days on average. As with enrolment, we do not observe any systematic differences between girls' and boys' attendance rates.

There is no evidence of an overall positive SCW effect on the reading fluency of girls in treatment communities compared to control communities. Girls in the treatment group can now read an average of 39 words per minute against a baseline of only 22 words per minute. But reading fluency improved just as much in the control group and a large part of the observed increase might be explained by girls growing older and progressing naturally in school. For example, international benchmarks for oral reading fluency (see Box 5) suggest that children should progress from reading 47 words per minute by the end of grade one, to reading 95 words by the end of grade 2[55]. In comparison to these benchmarks (although they were developed to assess children in comparatively high-performing education systems), an increase of 17 words per minute within the treatment group over two years appears small and may well be entirely due to natural progression.

Further analysis shows that the **effect of the SCW on oral reading has tended to focus around the middle of the score distribution (medium performers).** In particular, our EM demonstrates a significant positive effect on the median wpm score. At baseline, the median was 0 across SCW treatment areas, indicating than more than half the girls read 0 wpm. At midline, the median changed to 27 wpm, implying that half the girls could read at least 27 words-per-minute, compared to only 23 in control areas.

It is worth noting that although the EGRA oral reading subtask scores do not show any significant change over-and-above the control group across the SCW, the **letter sound and reading comprehension subtask scores have increased significantly more in the treatment than in the comparison group between baseline and midline**, by 2.4 out of 100 letters and 0.1 out of 5 questions more respectively. The invented word subtask does not show any significant change compared to the control group.

Similarly, numeracy scores have not improved significantly more in treatment than in control areas since baseline: EGMA scores increased from 30 per cent to 49 per cent in treatment areas compared with 30 per cent to 46 per cent in control areas. At the project level, ACTED (Afghanistan) and CfBT (Kenya) are the only project areas where improvements in literacy scores in the treatment group significantly exceeded those of the control group. When looking at EGMA scores, the Aga Khan Foundation (Afghanistan) project area is the only one which demonstrates a significant difference-in-difference.

Out-of-school girls are the only identified subgroup of marginalised girls that the SCW had a net positive effect on in terms of learning. Their oral reading EGRA scores increased by 12 words-per-minute more in treatment areas than in control areas, while their EGMA scores increased by 13 percentage points more in treatment areas than in control areas. For other subgroups the data shows no significant difference-in-difference in improvements between the treatment and control areas. While we can assert that the SCW has been effective in improving the learning of out-of-school girls, this claim cannot be made for other marginalised sub-groups. As opposed to literacy, our analysis of EM data shows that **the overall effect of SCW projects on EGMA scores has tended to focus on the lowest-performing girls**.

[55] See https://dibels.uoregon.edu/docs/marketplace/dibels/DIBELS-6Ed-Goals.pdf.

The activities which seem to have had the biggest impact on girls' learning are special tutoring and help with school work as well as teacher training. Those activities have impacted positively on both literacy and numeracy results. Again, control areas show similar patterns as treatment areas, which suggests that being exposed to these activities supports learning regardless of whether or not they are GEC-specific. Other activities such as the provision of bursaries and learning materials, or the supply of funds to a school were not found to have a direct and significant positive effect on literacy and numeracy scores.

Key lessons learned

- Gender analysis from the DRC, Ethiopia and Kenya project areas confirms our baseline findings that **in-school girls are not systematically disadvantaged compared with in-school boys of the same grades** in these areas. The gender gaps related to being-in-school and learning were relatively small at baseline. Girls in treatment and control areas showed greater improvements in literacy and numeracy scores than boys of the same age and school year. In other words, **girls are now overtaking boys across the SCW**. Although this jump might be due to the success of education interventions in both treatment and control areas, it raises the question of whether projects fully understood which educational barriers were gendered and which were universal, to what extent and where gender gaps existed in their treatment population when selecting target communities and designing interventions at the outset. These findings underscore that gender analysis should have been a core element of project design and M&E frameworks to ensure that projects truly promote equity in educational provision.
- SCW projects had a positive effect on the learning outcomes of out-of-school girls. They are the only subgroup to show a significant improvement over and above the change observed in control areas. This impact on out-of-school girls is a success and encouraging, given the lack of a detectable effect on the general population of treatment girls. It seems that relevant and effective mechanisms were put in place to reach and support these girls in particular. However, other groups of marginalised girls do not seem to have been reached and impacted to the same extent. **Girls who do not speak the official language of instruction have been shown to perform particularly poorly** at baseline and their literacy and numeracy scores increased the least among all subgroups. Those girls would require specific approaches to support their learning and it seems that the SCW has not been able to deliver these thus far. In fact, while 11 projects targeted out-of-school girls with particular interventions such as tutoring and alternative education, providing support to children with language difficulties was rarely mentioned in projects' midline evaluation reports and design documents. This emerges as an area that will require further attention from future education programmes.
- There is **little convergence between learning results measured through the analysis of EM data and the FM reanalysis of projects' data, but this is largely due to methodological differences**. According to the projects' own data, three out of five projects measuring their impact as a before-after intervention change hit their literacy target, whereas this was the case for four of the eight projects using control groups. In numeracy, none of the five projects using benchmarks instead of control groups hit their target, while two projects out of the seven using control groups had at least a high impact. But the differences in the sampling designs, target populations, learning assessments and methodologies used to assess impact make it impossible to draw any general conclusions of the aggregated impact of SCW projects based on their own data.
- The previous point also confirms the finding that **girls in treatment areas improved their reading and numeracy scores overall from baseline but this improvement was not significantly larger than the improvement achieved by girls in control areas at the window level**. It is consistent with the findings in Section 3.2 that target communities have been more exposed to education-related activities at midline than at baseline, but to a similar extent in treatment and control areas. This phenomenon is due to a mix of factors including: the insufficient size and/or scale of impact achieved by SCW projects; overcrowding of target communities with GEC and non-GEC actors; possibility of positive spill over effects, and poor sampling from the onset (the lists of sampling points provided by projects before baseline may not accurately reflect their spaces of intervention anymore, and some sampling points may have been too close together to prevent large spill over effects from treatment to control areas – see Annex F).

3.4 What has worked, why and with what effects?

This section explores the effects of SCW projects' activities on barriers to girls' education and assesses what activities have been the most effective, and in what contexts. Our analysis relies on EM's and projects' qualitative and quantitative data, as well as data provided by the FM:

- **Tables mapping what activities each project undertook to address a particular type of barrier** are informed by an intervention mapping conducted centrally by the FM for all SCW projects in 2015. The systematic categorisation of intervention types presented below was developed by the FM for this purpose. Since completing the mapping the FM has provided some specific updates that we have taken into account.
- Tables showing the **levels of barriers across project areas at midline** are based on the **EM reanalysis of projects' quantitative data**. Red cells indicate higher levels of a variable coded negatively (as an assumed "barrier" to education) while green cells indicate higher levels of variables expressing positive changes. The EM was not able to compare project data from baseline and midline because most projects did not merge their full household survey data from both waves.
- **Difference-in-difference analysis** has been carried out using EM baseline and midline household and school-visit surveys, as well as school-based assessment data in countries where this data was collected. All variables in the difference-in-difference tables are coded negatively, as barriers to education. A negative change of these variables must therefore be interpreted as the result of a potential positive effects of SCW projects' intervention (green cells). In contrast, a positive change suggests a negative effect of projects' activities (orange cells).
- The quantitative analysis is **complemented with data from the EM's qualitative IDIs and projects' midline evaluation reports**.
- **Expenditure data by project was extracted from the FM Value for Money analysis** (see footnote 58 for more details).We generally retained the higher-level budget categories as they had been assigned by the FM, but often aggregated more specific budget lines nested within overarching categories.

As far as possible, difference-in-difference results are presented as follows: (Treatment Baseline / Treatment Midline, Control Baseline / Control Midline, Difference-in-Difference indicator[56]). Tables showing full difference-in-difference results across all analytical variables from the EM HHS, SVS and SBA as well as from projects' household surveys are shown in Annex H.

Box 7: A brief caveat around the analysis of effectiveness

In Section 1.1.3 we have explained that SCW projects have taken a holistic approach to intervention, typically combining a range of different activities to address a variety of barriers to girls' education. This has been a characteristic feature of the SCW design from the start.

For the evaluation, however, this means that we cannot disentangle the effect of each individual intervention through statistical analysis. This would have required projects to allocate single interventions to clearly delineated groups or locations, whose outcomes we would then have compared systematically to those achieved by other interventions and groups who did not receive any intervention (i.e. the control group). This, however, was not a requirement of the GEC. Instead, a bundle of interventions is being delivered in each project area and projects have not recorded which specific interventions each individual girl beneficiary has received. Given that projects are reaching tens of thousands of girls through a combination of direct and indirect (e.g. community-based) interventions, such record keeping would have been extremely difficult to achieve, and was equally not required. Besides, it is extremely difficult to capture such information through household or school surveys with a sufficient level of precision and objectivity (this has been shown in Section 3.1 on exposure and reach).

Due to these design features it is impossible for us to disentangle the contribution that each individual intervention has made to changes in outcomes. For similar reasons, we do not produce any direct statistical analysis of the causal link between the change in barriers and the change in outcomes. Such an analysis would require building a model that virtually includes all variables that can affect the changes of a being-in-school or learning outcome. While we could in theory include in the same regression a series of variables harvested from

[56] The DID indicator will be followed by an asterisk if p-value is significant at the 10% level and by two asterisks if significant at 5%.

the different EM surveys, it is likely that this set of variables will not be enough to explain most of the observed change in outcomes. We did run such a multivariate analysis in our EM SCW Baseline Report but were able to explain only a very small proportion of the attendance and learning variables at the window level. Lots of factors, such as the institutional context, cannot be accurately captured through our surveys. Moreover, interpreting and building such models requires thorough contextual analysis and qualitative analysis of the process of change at the project level, which we do not have.

This section looks at changes in barriers that logically link to specific intervention types. Where we observe a significantly larger change in a particular barrier (e.g. affordability of schooling) in the treatment group compared to the control group we suggest that this *might* indicate that a relevant intervention carried out by the project (e.g. bursaries) has had some effect.

3.4.1 Poverty and livelihoods

All 14 SCW projects identified poverty and livelihood issues as key barriers to girls' education. This includes the affordability of school fees and learning materials (e.g. textbooks, or uniforms), as well as the gender division of labour within households that can make it difficult to attend school regularly and find time to study. For example, situations where girls spend a significant share of their day doing household chores, caring for other family members, or working outside the house.

Summary of baseline findings

At baseline, we found that poverty and livelihood issues play a crucial role in limiting girls' access to education and learning while in school. Girls from families with an unstable income and girls with significant domestic chores had a higher risk of being out of school. Qualitative evidence showed that difficulties in affording the costs of schooling was a key reason why girls dropped out of school, especially when they reached secondary school age. In some project areas, girls were being excluded from school for not having paid their school fees or because they did not have a uniform. Both the qualitative and quantitative evidence suggested that girls' household tasks, such as looking after younger siblings and fetching water had adverse effects on their learning.

Economic interventions addressing poverty and livelihood issues

As shown in Table 17, all SCW projects tackle poverty and livelihood issues as a core part of their theory of change. Eleven projects are providing in-kind support in the form of school kits, uniforms, textbooks, stationery or menstrual supplies. Eight projects also provide bursaries or scholarships to cover the cost of schooling. Five projects have established village or community saving loan schemes to help families afford the cost of schooling, and some have engaged families in income-generating activities (4 projects) or conditional cash transfer schemes (CfBT in Kenya).

Table 17: Number of SCW projects addressing poverty and livelihood issues

ECONOMIC	Total core	BRAC	AKF	Acted	IRC	STC	ChHp	WUSC	CfBT	STC	Plan	RI	CARE	WV	Camfd
		Afg	Afg	Afg	DRC	Eth	Eth	Ken	Ken	Moz	Sie	Som	Som	Zim	Z-T
Bursaries	8	✓			✓			✓		✓	✓	✓	✓		✓
Cash Transfers	1								✓						
Income-generating activities	4			✓			✓							✓	✓
In-kind support (school kits, menstrual sup.)	11	✓	✓	✦		✓	✓	✓	✓	✓	✓	✓	✓	✓	
Loans and savings	5				✓	✓	✓							✓	✓
Total	14	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Key: ✓ = The intervention is at the core of the project's intervention strategy. ✦ = The intervention is being delivered, but is not a core activity.

Source: Adapted from the FM's intervention mapping.

Since the GEC is an education programme rather than a livelihoods programme, projects cannot be expected to successfully alleviate severe poverty within three years. But the project interventions mentioned above can reduce the effects of poverty on girls' education, for example by helping families reallocate resources or bridge gaps in

ways that enable them to send their girls to school. We can also assess whether the burden of household chores and livelihood activities on girls has decreased since baseline, as a result of families prioritising education.

Midline evaluation findings

Project data suggests that across the SCW a large proportion of a girl's day is spent on household duties. Besides, a large percentage of caregivers report that this work is preventing girls from going to school and spending more time on school work (see Table 18). This phenomenon exists across both treatment and control areas for all projects that reported data against this indicator at midline.

Table 18: Midline project data on poverty and livelihoods issues

Midline barriers by project Project data		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
HHS															
Proportion of girl's time spent on household duties	T	11%			25%	57%	30%	26%	26%				43%	17%	
	C				27%	56%	30%		25%					21%	
Household chores prevent girl going to school	T	5%				73%	19%	100%	6%			18%	21%	4%	
	C					61%	19%		7%					5%	
Household chores reduce time spent on school work at home	T	14%				44%	35%	31%	11%				19%	7%	
	C					33%	33%		8%					8%	

Key: Darker shades of red indicate higher prevalence of the related barrier.

The highest figures are reported by Save the Children (Ethiopia), where more than half of a girl's day is allocated to household duties and where 53 per cent and 44 per cent of treatment households respectively report that chores prevent girls form going to school and that chores reduce time spent on school work at home. In contrast, less than 10 per cent of caregivers in World Vision (Zimbabwe) project areas report that household chores prevent girls going to school or reduces their time spent on school work.

Baseline to midline changes

At the window level there was no significant improvement in poverty and livelihood factors at midline in the treatment group, compared with the control group (see Table 19). In fact, the proportion of girls spending more than one hour a day **doing non-school work** increased from 44 per cent at baseline to 57 per cent at midline in the treatment group (and to a similar extent in the control group). This might reflect that the girls are now two years older and may have been given greater responsibilities in the household. In line with this, the share of households reporting that household chores stop the girl from attending school has increased from 17 per cent to 29 per cent.

The share of households in the treatment group that find it difficult to afford the cost of schooling has also increased from 44 per cent at baseline to 53 per cent at midline. This increase may equally be linked to the girls growing older and transitioning from primary to secondary school, secondary school fees being often higher than primary school fees.

Table 19: Difference-in-Difference in poverty and livelihoods issues

Difference-in-difference by project EM 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 Z-T
HHS															
Girl spends more than one hour a day on non-school work (HHS)	-3%	-10%	-18%	-2%	7%	-1%	0%	-25%	10%	12%	-16%	-7%	-7%	7%	
Household chores stop girl going to school (HHS)	-2%	-11%	-21%	-15%	-3%	-5%	-2%	18%	3%	-2%	-6%	0%	-3%	4%	
Difficult to afford for girl to go to school (HHS)	-3%				0%	-1%	-8%	-22%	4%	4%	8%	9%	4%	-6%	

Key: The difference-in-difference indicator is equal to: DID = (Midline Treatment – Midline Control) – (Baseline Treatment – Baseline Control). The way variables are coded implies that *negative* DID demonstrate a *decrease of barriers to education*, hence a "positive" effect of the SCW.
Light green = barrier decreased significantly at the 10% level. Light orange = barrier increased significantly at the 10% level.
Darker green = barrier decreased significantly at the 5% level. Darker orange = barrier increased significantly at the 5% level.
Grey = no figure or sample size smaller than 50 individuals per project area per wave.

At project level, two projects seem to have had a large and significant effect on the amount of **time girls spend doing household chores** and non-school work. The share of girls spending more than an hour per day on chores

increased “only” by 5%-points in the AKF (Afghanistan) treatment group, while increasing an additional 18%-points in the control group (45% / 50%, 42% / 65%, -18* %-p). Likewise, the share increased slightly from 42% to 47% in WUSC (Kenya) project areas, whereas it rose by an additional 25%-points in control areas (42% / 47%, 20% / 51%, -25** %-p). This suggests that AKF and WUSC have been able to prevent girls from taking on more responsibilities in the household as they grow older.

It is worth noting that 18 per cent more caregivers report that household chores stop girl going to school in WUSC treatment than in control areas. While this seems to contradict the positive effect observed on the time spent doing non-school work, this trend may also reflect the higher awareness of this barrier on girls’ enrolment and attendance. It may also be that caregivers having invested in girls’ schooling, become more reluctant to load them with household tasks that might stop them from doing well at school.

In Aga Khan Foundation project areas, our qualitative research found that girls’ spend less time doing housework in treatment than in control areas. Respondents suggest that this is largely due to community awareness-raising activities and engagement of community members in school governance. Although these did not belong to the main activities delivered by the AKF-led STAGES project, the latter seems to have had a positive effect on girls’ time spent doing non-school work:

> *“If a girl thinks about her future she should not leave her education. If she can’t finish school at least she should study 80 percent of school – and no girl has left school due to chores or other work at home.”* (IDI with Community Leader, Baghlan, AKF, Afghanistan)

More details on the effect of activities aimed at improving communities’ attitudes are presented in Section 3.4.5.

Box 8: World Vision Zimbabwe – the BEEP Bicycle scheme

World Vision established a Bicycle Education Empowerment Programme (BEEP), which is a community-based scheme to provide bicycles to students that live far away from school to help them enrol and more regularly attend school. The project’s midline evaluation report states that the scheme has reduced girls’ time poverty and tardiness in school and has established greater gender equality through girls’ access to bicycles.

Respondents in the EM’s qualitative interviews also spoke positively about BEEP. Some noted that it encouraged enrolment and consistent attendance in school, but it was not always clear whether this was because getting to school was now easier, or whether girls now made a greater effort to attend so that they could keep their bicycles. There is some qualitative evidence from the EM research, which suggests that the time girls save by cycling to school is actually spent on household chores rather than studying.

We find that household chores and difficulties in affording the cost of schooling have generally become greater as the girls have grown older, which is in line with the qualitative evidence collected at baseline.

Difficulties in affording the cost of schooling were reported by 91 per cent of households in World Vision (Zimbabwe) project areas at baseline, compared with 90 per cent at midline. In the control group, the share increased, resulting in a significant difference-in-difference (91% / 90%, 88% / 94%, -6* %-p). World Vision invested about 40 per cent of their budget in forming Village Savings and Loans (VSL) groups and in providing in-kind support such as sanitary pads. Over the past couple of years, the country’s macroeconomic situation has deteriorated. Zimbabwe has faced a severe and prolonged drought affecting mostly the poorest households. In the first quarter of 2016, about 1.5 million people were reported to be food insecure[57]. This was emphasised by the project’s evaluation midline evaluation report and our qualitative analysis which found that World Vision’s VSL groups helped improving households’ livelihoods and enabled families to afford the cost of schooling. However, limitations to this intervention were also reported like the failure of some of the VSL groups as a result of the severe drought and the subsequent need of local populations to access savings.

In WUSC (Kenya) treatment areas, the share of families struggling to afford the cost of schooling increased from 27 per cent to 44 per cent, but it rose an additional 22 percentage-points in the control group (27% / 44%, 28% / 67%, -22 %-p). Even though this difference-in-difference is not statistically significant, it is large and suggests that WUSC have been able to remedy some of the increase in costs that arises as girls grow older through the provision of scholarships covering tuition fees and education supplies to secondary school girls. Our qualitative research found

[57] http://reliefweb.int/sites/reliefweb.int/files/resources/WFP%20Zimbabwe%20Situation%20Report%20%233%2C%2010%20January%202016.pdf

that the cost of schooling is a bigger barrier to education in control than in treatment areas. In treatment areas, many respondents mentioned that the provision of sanitary pads and scholarships have had a positive effect on households' ability to afford school:

> *"Even the sanitary pads have been provided to prevent the girls from having problems when in menstruation, so some parents may not afford the cost of the sanitary pads but when the NGOs are giving out, they are very much appreciative."* (IDI with community leader in Garissa, WUSC, Kenya)

Budget expenditure and effectiveness

Table 20 below shows what percentage of their budget projects spent on economic interventions to address poverty and livelihood issues.[58] On average, SCW projects are spending 22 per cent of their budget on economic activities, but there is significant variation across the portfolio. For instance, Camfed (Tanzania and Zimbabwe) and ChildHope (Ethiopia) both spent about 42 per cent of their budget whereas CARE (Somalia) only spent 2 per cent.

Out of SCW projects, IRC (DRC) spent the most on economic interventions (54 per cent of their budget). However, they have not had any significant effect on economic-related barriers. IRC runs economic and social empowerment trainings with their target communities and provides scholarships to primary school students. Data from our household survey show that the proportion of caregivers reporting that school is difficult to afford has increased by 5 percentage-points in both treatment and control areas (75% / 80%, 70% / 75%, 0 %-p). In the meantime, the proportion of girls spending more than one hour a day doing non-school work has increased even further from baseline to midline (37% / 55%, 38% / 48%, 7 %-p). Our qualitative research found that although IRC's economic interventions were generally well perceived by respondents, this has not led to a substantial decrease in the cost of schooling. Some respondents also mentioned that loans and savings groups did not work partly because the mechanisms were too complicated, or because some borrowed money to purchase animals but there were no market to sell them. The project's midline evaluation report includes lessons learned about the need to closely monitor these groups to avoid lending money to groups who will not easily be able to invest it. The report also noted that the payment of scholarships was delayed, resulting in students dropping out or being chased away[59].

Table 20: Percentage of the project budget spent on economic interventions

ECONOMIC	Avg. spent	BRAC	AKF	Acted	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
Bursaries	7%	6%			41%			15%		8%	N/A	13%	0%		38%
Cash Transfers	2%								9%	8%	N/A				
Income-generating activities	1%						7%				N/A				
In-kind support (school kits, menstrual supplies)	10%	12%	8%				35%	14%	9%	7%	N/A	12%	1%	14%	
Loans and savings	3%				12%	11%					N/A			25%	4%
Total share in budget expenditure	25%	17%	8%	0%	54%	11%	42%	30%	18%	23%	N/A	25%	2%	39%	42%

Key: The blue colour coding indicates which specific intervention subcategories a project spends the most money on. The darker the shade of blue, the higher the proportion of overall budget invested in a specific intervention type.
The violet colour coding compares horizontally across projects, indicating which project invests the most in the overall intervention category. The darker the colour shading, the more money was invested relative to the spent of other projects.
Note (*): Numbers shown for IRC and STC Ethiopia are derived from Year 3 expenditure, as opposed to Years 1 & 2 for other projects.
Source: Budgetary information and assigned categories were extracted from the FM's Value for Money analysis, and adapted if needed.

In project areas in which poverty and livelihood factors improved significantly, we find that both WUSC (Kenya) and World Vision (Zimbabwe) invested relatively large shares of their budget in economic interventions (30% and 39%

[58] All projects with the exception of Plan (Sierra Leone) provided this costing data to the Fund Manager who then categorised different budget lines according to the GEC intervention categories. This data has been sourced from the Value for Money analysis completed by the FM in August 2015 and as such does not represent the full two-year expenditure profile of the projects, which may have been subject to change throughout the implementation period depending on project spend profiles. Amounts of budget spent have been aggregated from different activities and categorised across outputs by the FM. Expenditure data for IRC (DRC) and Save the Children (Ethiopia) was only integrated to the Value for Money analysis in October 2016. As a result, the budget of those two projects reflect their expenditure for Year 3 only, while for other projects our analysis is based on combined Year 1 and Year 2 data dating back from August 2015.
[59] More details are given in the EM Thematic Research Report "Narrow Windows, Revolving Doors" available here: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/520249/Narrow-Windows-Revolving-Doors-What-affects-adolescent-girls_-ability-to-stay-in-school-Research-Report-March-2016.pdf

respectively), whereas AKF (Afghanistan) invested relatively little (8%). All three projects provided in-kind support in the form of school kits, uniforms, books or sanitary supplies. In addition, WUSC provided bursaries, and World Vision ran a loans and savings scheme. As mentioned above, our findings suggest that AKF and WUSC have been able to prevent girls from taking on more responsibilities in the household as they grow older – WUSC run sensitisation trainings for parents while AKF worked with school actors (management councils, shuras, teachers) to help them follow up on girls' absenteeism and raise parents' awareness on the importance of education versus domestic chores or livelihood activities.

Except IRC, the three projects that invested the most in economic interventions are those showing the strongest effects in improving school affordability. This tends to confirm the findings from many project midline evaluation reports that highlighted the effectiveness of cash and in-kind support in removing the key barrier of affordability of schooling.

3.4.2 Teaching

Eight out of the initial 15 SCW projects mentioned issues with the quality of teaching as a key barrier to girls' education. Issues included poor teaching practice and pedagogy, frequent teacher absence, lack of qualified teachers and inappropriate curricula.

Summary of baseline findings

Our baseline analysis found a significant association between parents' dissatisfaction with teaching and teacher absence on the one hand, and poor student attendance and reading fluency on the other hand. In EM IDIs, respondents across the SCW reported a general lack of teachers, inadequately qualified teachers and a frequent absence of available teachers as key barriers. In the EM's school-based assessment, 77 per cent of schools reported a lack of teachers, and only one in four available teachers had a teaching certificate.

The lack of teachers was often linked to overcrowded classrooms and high student-teacher ratios that made it difficult for teachers to tend to individual students. This was supported by the quantitative evidence, which showed that girls in large classes achieve lower average reading scores after controlling for other factors.

Interventions to improve the quality of teaching

Table 21 shows the approaches taken by SCW projects to improve the quality of teaching in intervention schools. A large number of projects are listed as training teachers in gender sensitive pedagogy as a key element of their theory of change. However, the actual budget spent on gender responsive pedagogy was comparatively small (see Table 24), which may be because it is often integrated into more general teacher training. Half of all SCW projects train teachers in teaching literacy and numeracy, and six are building teachers' pedagogical skills more generally.

Despite the general shortage of teachers in many SCW project areas, only three projects placed the recruitment of teachers at the core of their theory of change. These include BRAC and AKF (Afghanistan) who recruited teachers for newly established community based schools, and WUSC (Kenya) working in Dadaab and Kakuma refugee camps. Many projects working in established government schools would have struggled to recruit new teachers whose training and hiring is typically the government's responsibility. The lack of qualified teachers is indeed a systemic issue that projects can address in some ways (e.g. by suggesting changes to the teacher training curriculum) but cannot fully address without aligned government action.

Table 21: Number of SCW projects addressing quality of teaching

	Total core	BRAC	AKF	Acted	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
Formal pre-service teacher training	1		✦								✦		✓		
Gender responsive pedagogy	9		✓	✓	✓	✓	✦	✓	✓	✦	✓	✓	✓		
Inclusive classroom strategies	3	✓	✓								✓				
Literacy and numeracy	7	✓	✓		✓	✓			✓	✓			✦	✓	
Peer support and mentoring	4		✓						✓	✓	✓	✦	✦		
Recruiting (female) teachers	3	✓	✓					✓					✦		

Skills training	6	✓	✓	✦			✓			✓	✓				✓
Total	14	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Key: ✓ = The intervention is at the core of the project's intervention strategy. ✦ = The intervention is being used, but is not a core activity.

Source: Adapted from the FM's intervention mapping, accounting for updates provided in January 2017.

Midline findings

The proportion of households that have reported an increase in the number of teachers and in the quality of teaching since baseline varies widely across contexts (see Table 22). Caregivers have reported improvements in the quality of teaching more often than an increase in the number of teachers. The greatest changes are in IRC (DRC) where almost all caregivers reported an increase in the number of teachers since baseline. They were more than half to do so in CfBT (Kenya) but less than 15 per cent in BRAC (Afg), Relief International and Care (Somalia) areas. However, data suggests that control areas experienced similar increases in the number of teachers. The same trend is observed regarding the quality of teaching. This could be explained by improvements in public education provision, the implementation of other interventions in treatment areas, or by SCW activities spilling over into control areas. Increases in the number of teachers seem generally correlated with increases in the quality of teaching, except in BRAC (Afghanistan) and RI (Somalia) where quality has improved but not quantity, and in IRC (DRC) where the opposite pattern is observed.

Table 22: Midline project data on barriers related to teaching

Midline barriers by project Project data		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
HHS															
Increase in the number of teachers since baseline	T	14%			97%	33%	37%		54%			0%	14%	40%	
	C				92%	30%	35%		61%					44%	
Increase in the quality of teaching since baseline	T	62%			33%	36%	49%		67%			40%	26%	53%	
	C				27%	39%	41%		75%					52%	

Key: Darker shades of green indicates higher level proportion of caregivers reporting a positive change since baseline.

Baseline to midline changes

Table 23 shows whether barriers relating to the quality of teaching have improved in treatment areas, compared with control areas. Drawing on data from the EM household survey we assess parental satisfaction with the quality of teaching at the girls' school and the frequency of teachers being absent. We use data from the school visit survey to measure the share of female teachers in SCW schools and the incidence of classroom overcrowding. We also use data from the school-based assessment to assess the share of teachers without a teaching certificate.

It is worth noting that parental perceptions and awareness of changes in the girls' school may be partly subjective and affected by a number of factors other than the intervention itself. Some projects may simply communicate their activities more effectively than others. Besides, concepts such as the quality of teaching may highly depend on people's expectation, and not be accurately captured through survey-based measures. The analysis presented below needs to be interpreted with those caveats in mind.

At the window level, we find a relatively small but significant effect on **parents' satisfaction with the quality of teaching** at the girl's school. At baseline, 26 per cent of parents in the treatment group were dissatisfied with the quality of teaching. At midline, the share had dropped to 23per cent (2% / 23%, 28% / 30%, -5* %-p).

At the project level, we measure a relatively large difference-in-difference in the three Afghan project areas, as well as in Save the Children (Ethiopia) and Relief International (Somalia) project areas. Only in Relief International (Somalia) and AKF (Afghanistan) project areas is the effect statistically significant. In AKF project areas, 37% of parents in treatment areas were dissatisfied with the quality of teaching at baseline, compared with only 16% at midline, while dissatisfaction increased in the control group (37% / 16%, 21% / 35%, -35** %-p).

In Relief International project areas, parents were *less* satisfied, on average, with the quality of teaching than they had been at baseline. The share of those stating that teaching was not satisfactory increased from 10 to 14% in the treatment group. However, this increase in dissatisfaction was much more pronounced in the control group (8% to 28%) so that the significant difference of -16**%-p between the treatment and control group points towards a positive development.

With regards to **teacher absence**, we find no positive effect of the SCW at window level. On the contrary, teacher absenteeism was reported significantly more often in treatment areas than in control areas, which might be due to respondents in treatment areas becoming more aware of this phenomenon. This is observed most strongly in WUSC (Kenya) project areas where the share of parents reporting frequent absences of the girl's teacher rose from 8 per cent to 25 per cent. In the control group, the share remained stable around 20 per cent (8% / 25%, 20% / 21%, +16* %-p). Rather than suggesting that projects have made teacher absenteeism worse, it is possible that treatment communities have become more aware and outspoken about teacher absenteeism as a problem, as a result of projects' awareness raising activities.

Our qualitative findings, the project's midline evaluation report and other secondary documents suggest that the general worsening of teachers' absences may be the result of several factors including the threatening security situation in Dadaab and Kakuma refugee camps and its host communities, which has led many teachers to leave the area, and the national teachers' strike that has been ongoing since the beginning of the project. Moreover, teachers appointed to remote communities might be more prone to demotivation due to the lower wages earned compared to the rest of the country[60], which could eventually lead to more frequent absences.

School-based data confirms that the teacher shortage intensified in WUSC project areas between baseline and midline. Only in CfBT (Kenya) project areas do activities seem to have had a strong positive effect on the availability of teachers with shortages reported in only 49 per cent of school in treatment areas at midline compared to all of them at baseline.

Table 23: Difference-in-Difference in barriers relating to teaching

Difference-in-difference by project EM 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 Z-T
HHS															
Caregiver thinks teaching is not satisfactory at girl's school	-5%	-10%	-35%	-15%	7%	-11%	-5%	9%	1%	-6%	4%	-16%	4%	1%	
Teacher is absent at least a few times a month	5%	11%	-10%	-1%	2%	-3%	1%	16%	5%	9%	-3%	11%	9%	1%	
SVS															
Girl's main teacher is a male	1%				7%	1%	-6%	-4%	5%	-4%		2%	5%	4%	
There are 40 children or more in girl's class	0%					13%	-7%	-8%	-8%	6%		24%	24%	-9%	
Average number of hours taught per day	-0.3					0.1	0.2	-0.8	-0.3	-0.4		0.8	-1.1	0.2	
SBA															
Teacher has no teaching certificate	2%				3%	-6%	16%	-8%	11%						
Shortage of teachers in the school	-8%				-11%	-1%	17%	17%	-52%						

Key: The difference-in-difference indicator is equal to: DID = (Midline Treatment – Midline Control) – (Baseline Treatment – Baseline Control). The way variables are coded implies that *negative* DID demonstrate a *decrease of barriers to education*, hence a "positive" effect of the SCW.
Light green = barrier decreased significantly at the 10% level. Light orange = barrier increased significantly at the 10% level.
Darker green = barrier decreased significantly at the 5% level. Darker orange = barrier increased significantly at the 5% level.
Grey = no figure or sample size smaller than 50 individuals per project area per wave.

We find no significant change in the ratio of male to female teachers at the window level. We also did not find any effect on the average class size, the average number of hours taught per day, the proportion of teachers holding a teaching certificate or the availability of teachers at school. This is not surprising as most of these dimensions largely depend on institutional factors that we do not expect projects' activities to impact on over such a short period of time. It appears that classroom overcrowding has become more prevalent in project areas where enrolment was very low at baseline and where activities focused on getting girls into school (Relief International and Care (Somalia).

[60] https://www.cigionline.org/sites/default/files/cigi_pb_47.pdf

Budget expenditure and effectiveness

Across the SCW, projects are spending on average 18 per cent of their budget on teacher training interventions (see Table 24). There is a lot of variation between projects, with BRAC (Afghanistan) spending over half of their budget on teacher training, whereas no spend is shown for World Vision (Zimbabwe). The latter only launched their "Happy Reader" literacy and numeracy training for teachers around the time when midline was conducted, and this is not yet reflected in the budgetary data displayed below.

ACTED is also shown to have spent almost nothing on formal teacher training. However, they did recruit and train teachers specifically for their Youth Development Centres and Community Literacy Classes and have accounted for these costs in the category "Extracurricular activities and non-formal education" (see Section 3.4.8).

Table 24: Percentage of the project budget spent on teacher training interventions

	Avg. spent	BRAC	AKF	Acted	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
Formal pre-service teacher training	3%	23%	7%								N/A		7%		
Gender responsive pedagogy	1%		1%			1%	4%		2%		N/A				
Inclusive classroom strategies	0%										N/A				
Peer support and mentoring	0%		5%								N/A				
Skills training (including in teaching literacy/numeracy)	6%	3%	7%	0.1%	11%				22%	9%	N/A	15%			13%
Teachers recruited (e.g. female teachers)	5%	27%	20%					11%			N/A		3%		
Total investment	18%	53%	41%	0%	11%	1%	4%	11%	24%	9%	N/A	15%	10%	0%	13%

While satisfaction with teaching deteriorated significantly less in Relief International (Somalia) treatment areas than in control areas, AKF (Afghanistan) is the only project showing a significant difference-in-difference *and* an actual improvement in caregivers' satisfaction with the quality of teaching. AKF invested 41 per cent of their budget in the domain of teaching. Only BRAC (Afghanistan) invested more. AKF's spending is spread across a range of activities, such as recruiting teachers, formal and pre-service training, mentoring, and skills training.

BRAC (Afghanistan) and CfBT (Kenya) who also spent considerable amounts on teacher training (i.e. 53% and 24% of their budgets, respectively) did not show a net effect in the EM data. However, these two projects show the highest reported increase in the quality of teaching in their own project midline data (62% and 67% respectively, see Table 22). IRC (DRC) carried out teacher training and provided scholarships to students, which indirectly guaranteed the salary of teachers. Although the evidence from the project's midline report and the EM quantitative data is partly inconclusive about the effect of these interventions on teachers' absence, this could have made teachers more committed and present at schools. This has been confirmed by community leaders through IDIs. This perception may explain why IRC project's areas show such a large positive change in the increase in the number of teachers since baseline, according to the project's own data (see Table 22).

Across SCW projects, BRAC (Afghanistan) and AKF (Afghanistan) invested the most budget in recruiting teachers, but only AKF managed to lower perceptions of teachers' absenteeism, although not significantly. These two projects have mostly recruited teachers for their newly built community-based schools, which explains why they did not impact on teachers' absenteeism overall. CfBT (Kenya) seem to have been very successful in reducing shortages of teachers in school, although they did not recruit any teachers. Other types of activities such as mentoring or training may have acted as incentives for teachers to be more present in school, although high turnover has been reported by CfBT as a persistent issue in schools in urban slums.

Box 9: Example of teacher training and recruitment in AKF (Afghanistan) project areas

Hiring and training teachers and mentors is an important intervention in AKF's project STAGES and represents the largest budgeted area for the project. The project focuses on establishing Community-Based Education (CBE) at primary and lower secondary school levels and Accelerated Learning Programme (ALP) classes in communities where there are no government schools.

AKF carry out a range of activities:

- In-service training and mentoring on learner-centred and gender-friendly methodologies for teachers in Early Childhood Development, Community Based Education, Accelerated Learning Programme and formal schools;
- Apprenticeship teacher training and grants for young women from the target communities to enable them to attend teacher training college; and
- Training for staff of teacher training colleges, as well as provincial and district education departments to help sustain any benefits achieved by the project.

In the EM's qualitative interviews, the teacher training provided by AKF was mentioned frequently. Respondents reported that teaching practices had improved, and that teachers who completed AKF training appeared more 'professional' than their counterparts:

> *It has changed the people's views because before, teachers had limited knowledge of teaching, but now they have become more professional to teach our children based on the modern methodologies of teaching implemented in the world.*
>
> (IDI with a community leader, Molayan, AKF treatment area).

Teachers were reported to use more interactive teaching methods and encourage more discussion in class, and teachers themselves said that they were planning their lessons differently as a result of the training and felt that their teaching had improved. Respondents stressed that better teaching had led to children being more enthusiastic about school, and more willing to attend regularly. This evidence supports the quantitative findings presented above and suggests that teacher training in AKF project areas has been particularly effective.

3.4.3 School infrastructure and facilities

Many projects assumed that a lack of adequately equipped classrooms hampers learning in school. In their theories of change, seven projects mentioned inadequate school and classroom facilities as a barrier to girls' education, three mentioned the poor availability of schools in general, and another three explicitly mentioned inadequate sanitation facilities. Classrooms may lack solid floors or ceilings to protect children from the weather; access to electricity and water; sufficient desks or seating for children. Toilets may not be lockable, or there may only be common toilets for boys and girls.

Summary of EM baseline findings

Our baseline analysis of barriers to education assessed a variety of school-based factors. In the IDIs, respondents frequently mentioned difficulties with learning in school due to a **lack of proper school buildings** (including steady floors and roofs), poor access to a safe water source, and a lack of benches and chairs in the classroom. Our quantitative analysis did not find any significant relationship between these factors and girls' learning, but girls were found to achieve lower reading scores in classrooms without electricity – an issue that affected 91 per cent of classrooms. While it is plausible that a lack of electricity could hinder learning, a lack of electricity may also correlate with unobserved factors such as the resources available to the school or the community in general. The evidence was inconclusive with respect to the importance of separate and lockable toilets for girls' learning outcomes.

With regards to the **accessibility of schools**, a lack of schools and long journeys to available schools were mentioned in some IDIs, but were not a frequently reported barrier. There was a marginally significant association between the distance to the nearest school and enrolment, suggesting that girls in very remote areas may face the greatest obstacles to enrolment.

Interventions to improve school infrastructure and facilities

As shown in Table 25, all projects except IRC (DRC), place the improvement of school infrastructure and available resources at the core of their theory of change. All three Afghan projects have built new community-based schools in remote communities, and several other projects have either built new school facilities or improved existing ones. Many projects are providing textbooks and learning materials to schools, and five have improved WASH facilities.

Table 25: Number of SCW projects addressing school infrastructure and facilities

INFRASTRUCTURE & RESOURCES	Total core	BRAC Afg	AKF Afg	Acted Afg	IRC DRC	STC Eth	ChHp Eth	WUSC Ken	CfBT Ken	STC Moz	Plan Sie	RI Som	CARE Som	WV Zim	Camf Z-T
School and classroom building/ improvement	7	✓	✓	✓		✓	✓	✓		✦			✓		
Textbooks & Learning materials	8	✦	✓	✦	✦	✓		✓	✓	✓	✓	✦	✦	✓	✓
Toilets & WASH facilities	5		✦	✓		✓	✦	✓				✓	✓	✦	
Total	13	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Key: ✓ = The intervention is at the core of the project's intervention strategy. ✦ = The intervention is being used, but is not a core activity.
Source: Adapted from the FM's intervention mapping.

Midline findings

The proportion of caregivers that have reported positive changes to the school infrastructure and facilities varies widely across project areas with CfBT (Kenya) and to a lesser extent ChildHope (Ethiopia) reporting the highest proportion of positive changes (see

Table 26). However, change is similar across treatment and control areas for most projects, except for IRC (DRC), Save the Children (Ethiopia) and World Vision (Zimbabwe) where slightly more caregivers reported positive changes in treatment communities than in control communities on average.

Table 26: Midline project data on barriers related to school infrastructure and facilities

Midline barriers by project Project data		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
HHS															
Positive change to school facilities since baseline	T	17%			35%	27%	39%		55%			0%	20%	57%	
	C				23%	27%	26%		52%					61%	
Positive change to the number of schools since baseline	T	16%			15%	32%	19%		41%			24%	5%	14%	
	C				15%	30%	16%		38%					8%	
Positive change to the number of classrooms since baseline	T	20%			16%	31%	45%		38%			0%		26%	
	C				11%	29%	33%		46%					30%	
Positive change to the quality of classrooms since baseline	T	24%				35%	43%		50%			31%	30%	36%	
	C					31%	33%		48%					30%	
Positive change to the number or quality of textbooks since baseline	T	58%				36%	19%		56%			36%	18%	20%	
	C					37%	21%		54%					16%	

Key: Darker shades of green indicates higher level proportion of caregivers reporting a positive change since baseline.

Baseline to midline changes

Table 27 shows the difference-in-difference in key variables relating to school infrastructure and resources. These include parents' satisfaction with the classrooms, toilets, and textbooks at the girl's school as reported in the EM's household survey, and the share of toilets without access to electricity or water, or without indoor toilets, as reported in the EM's school visit survey.

At the window level, we find no significant effect on the quality and availability of school infrastructure and resources, although caregivers' satisfaction and provision of equipment show slight overall improvements. We do, however, measure effects in a number of project areas. In AKF (Afghanistan) treatment areas, parental satisfaction with the quality of school facilities has improved significantly. The share of parents who were dissatisfied with the quality of classrooms decreased from 35 per cent to 10 per cent, dissatisfaction with toilets decreased from 54 per cent to 19 per cent and dissatisfaction with textbooks decreased from 51 per cent to 25 per cent. In control areas, levels of satisfaction either remained the same or worsened after baseline. We see similar patterns in parental satisfaction with classrooms and toilets in ACTED (Afghanistan) project areas, where we also find a large and significant effect on these barriers.

In ChildHope (Ethiopia) and CfBT (Kenya) project areas, the proportion of schools having a book corner has increased by a large amount, as well as the proportion of lockable toilets for girls in schools (although only

ChildHope show a significant effect). In WUSC (Kenya) project areas, many more schools in the treatment group have established a book corner since baseline, compared with schools in control areas. The same is true for the availability of audio-visual equipment.

Box 10: AKF (Afghanistan) school infrastructure improvements

AKF interventions to improve the school infrastructure and resources were mentioned several times in the EM qualitative research, mostly in treatment areas. Respondents were happy about the schools being built:

> *I am personally very happy for that school, children of our village would be able to attend this school and get their education.*
>
> (IDI with a Community Leader, Kandahar).

Respondents welcomed the connection of schools to the electric grid, as well as the provision of furniture and equipment. In some cases, it seems that communities had asked for chairs but received carpets instead, which was sometimes positively perceived (*'still better than sitting on the dusty floors'*) and sometimes not. A few respondents also reported tree planting and the building of school playgrounds in intervention communities.

IDI respondents brought up school infrastructure and resources more frequently at midline than at baseline. This might reflect a greater general awareness resulting from AKF's activities in this area, as well as a rise in community expectations, following initial improvements. The EM qualitative analysis further suggests that the provision of schooling closer to home contributed to a rise in community support in favour of girls and young children attending school. Some children have even been encouraged by their families to drop out of madrasas to enrol in newly built schools.

Interestingly, many respondents refer to AKF and the Education Directorate as providing these supplies and jointly responding to the needs which shows that the project is working closely with local authorities and communities.

In IRC (DRC) and CARE (Somalia) project areas, we find large and significant inverse effects. In IRC project areas, levels of dissatisfaction with classrooms, toilets and textbooks changed little in the treatment group, whereas caregivers in the control group were much more satisfied. This pattern is relatively consistent across all variables related to caregiver satisfaction from the household survey (see Table 27). In CARE (Somalia) treatment areas, levels of dissatisfaction increased, whereas they either decreased or remained stable in the control group.

IRC is the only project not investing in school infrastructure and resources as part of their intervention strategy (see Table 27). The project supports school committees in developing their own school improvement plans, most of which are reported to include improvements to the school infrastructure and equipment. However, in their midline evaluation report IRC explains that some of these plans have not yet been implemented:

> *"A few COPAs did not yet receive money, and others complained that the school director managed the money, and they were not associated in the process, even though they were trained and they themselves worked out the school-related needs to be addressed."* (IRC midline evaluation report, p. 20).

The EM midline analysis suggests that these efforts have not (yet) resulted in any tangible results on caregivers' perceptions.

Table 27: Difference-in-Difference in barriers relating to school infrastructure and resources

Difference-in-difference by project EM 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 Z-T
HHS															
Classrooms not satisfactory at girl's school	-5%	-5%	-29%	-25%	28%	-12%	-14%	-10%	-8%	-1%	-11%	-7%	17%	6%	
Toilets not satisfactory at girl's school	-2%	4%	-38%	-25%	28%	4%	-2%	-12%	1%	-2%	7%	-13%	19%	6%	
Textbooks not satisfactory at girl's school	-1%	8%	-25%	10%	22%	2%	-11%	-1%	-17%	1%	-1%	-27%	16%	2%	
SVS															
Children do not use textbooks in class	0%				6%	-4%	0%	2%	1%	0%		2%	-1%	-1%	
Children unable to take text books home	0%				2%		0%	1%	-3%	0%		-6%	-1%	2%	
School doesn't have electricity all day	7%					7%	13%	28%	8%	4%		12%	-3%	-8%	
School doesn't have access to water	0%					11%	-1%	-2%	-1%	-26%		22%	23%	-8%	

School doesn't have a roof	4%					-6%	3%	15%	7%	-5%		-12%	15%	1%	
School doesn't have indoor toilets	1%					-22%	11%	1%	3%	8%		-12%	-8%	-9%	
SBA															
% of girls who don't have writing materials	-3%				-12%	-8%	2%	5%	-3%						
% of girls who don't have a bench or desk	-0%				4%	-26%	2%	14%	4%						
School has no library or book corner	-12%				22%	12%	-39%	-15%	-34%						
School has no audio-visual equipment	-3%				5%	-3%	8%	-22%	-8%						
Girl's toilets are not lockable	-8%				30%	-22%	-30%	-12%	-7%						

Key: The difference-in-difference indicator is equal to: DID = (Midline Treatment – Midline Control) – (Baseline Treatment – Baseline Control). The way variables are coded implies that *negative* DID demonstrate a *decrease of barriers to education*, hence a "positive" effect of the SCW.
Light green = barrier decreased significantly at the 10% level. Light orange = barrier increased significantly at the 10% level.
Darker green = barrier decreased significantly at the 5% level. Darker orange = barrier increased significantly at the 5% level.
Grey = no figure or sample size smaller than 50 individuals per project area per wave.

Budget expenditure and effectiveness

SCW projects are spending on average 23 per cent of their budget on improving the school infrastructure and resources (see Table 28). Again, there is significant variation between projects, with Save the Children (Ethiopia) spending 82 per cent of their budget in this area, ACTED (Afghanistan), WUSC (Kenya), and Care (Somalia) spending close to half of their budget, and others such as CfBT (Kenya) or Save the Children (Mozambique) spending only a small fraction (less than 5%). On average, projects are spending most of this budget on school and classroom building or improvements, a bit less on the provision of textbooks and learning materials, and relatively little on the improvement of toilets and wash facilities.

If we compare expenditure with the effects that projects have had on infrastructure barriers, we find a mixed and somewhat inconclusive picture. Save the Children (Ethiopia) invested heavily in building new Alternative Basic Education (ABE) schools and teachers' residences, but this does not translate into any observable effect in the quality of schools' infrastructure. This may be due to newly built schools not having better or more satisfying infrastructures than already existing schools. Save the Children also spent 21 per cent of their budget printing textbook teacher manuals but the effect of this intervention is not directly captured by our data.

ACTED (Afghanistan) invested a lot in school and classroom building, spending 45 per cent of their budget on this activity alone, and our analysis shows significant improvements in the satisfaction with both classrooms and toilets. ChildHope (Ethiopia) spent 31 per cent of their budget on school infrastructure and resources, with 17 per cent assigned to improving WASH facilities. Our school-based assessment data shows a large although not significant increase in the proportion of schools equipped with lockable toilets in ChildHope's project areas, and we also find an increase in parental satisfaction with the quality of classrooms.

In contrast, CARE (Somalia) equally invested 36 per cent share of their budget on school and classroom building but has not shown any positive effects. In fact, it seems that school conditions have improved a lot more in control areas than in treatment areas. Many respondents from our qualitative research, positively assess the impact of CARE's interventions, but highlight their insufficient scale to meet local needs. For example, newly built classrooms have been described as already overcrowded and lacking toilets.

Likewise, WUSC (Kenya) invested 32 per cent of their budget in the provision of textbooks and learning materials, and aimed at providing each student with three textbooks. While the project's midline evaluation report and our qualitative research suggest improvements in the availability and quality of teaching and learning materials as a result of interventions, our quantitative findings do not show any significant improvement in textbook use and satisfaction, despite an increase in the prevalence of book corners in schools.

In AKF (Afghanistan) project areas, we find some significant improvements in satisfaction with classrooms, toilets and textbooks despite the fact that the project invested comparatively little (10%) in school infrastructure and resources.

Table 28: Percentage of the project budget spent on school infrastructure and resources

INFRASTRUCTURE & RESOURCES	Avg. spent	BRAC	AKF	Acted	IRC	STC	ChHp	WUSC	CfBT	STC	Plan	RI	CARE	WV	Camfd
		Afg	Afg	Afg	DRC	Eth	Eth	Ken	Ken	Moz	Sie	Som	Som	Zim	Z-T
School and classroom building/ improvement	12%	6%	5%	45%	4%	61%	9%	14%	1%	2%	N/A	15%	36%		

Textbooks & Learning materials	8%	3%	4%	2%	6%	21%	5%	32%	1%	1%	N/A	11%	1%		26%
Toilets & WASH facilities	3%						17%		2%		N/A	4%	3%	8%	
Total investment	23%	9%	10%	47%	10%	82%	31%	46%	4%	3%	N/A	29%	41%	8%	26%

Source: Budgetary information and assigned categories were extracted from the FM's Value for Money analysis, and adapted if needed.

3.4.4 Interventions to improve school governance

Only two projects explicitly stated in their theories of change that poor school governance was a barrier to girls' education. This included a lack of planning and accountability, as well as poor structures for joint decision-making between teachers, parents, and other community stakeholders through parent-teacher or school management committees.

Summary of EM baseline findings

We did not examine issues around school governance at baseline.

Interventions that address school governance and management

Over the two years since our baseline, school governance has emerged as a more prominent issue and all projects, with the exception of Plan (Sierra Leone), are working in this area as part of their core intervention strategies. Ten projects work with School Management Committees (SMCs), Parent-teacher Associations (PTAs) and other stakeholders to improve school governance and management, and joint decision-making about school development. Nine projects work with local or national educational authorities. This is an important activity area to ensure that SCW interventions align with government priorities and policies, and to raise the chances of activities and initiatives being taken on or even scaled up by government after GEC funding end. As discussed earlier, the three Afghan projects stand out for having established community-based education facilities.

Table 29: Number of SCW projects addressing school management and governance

SCHOOL MANAGEMENT & GOVERNANCE	Total core	BRAC	AKF	Acted	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
Community and private schooling provision	3	✓	✓	✓											
Technology for school management	1								✦						✓
Working with local or national education authorities	9	✦	✓	✓		✓	✓		✓	✓		✓	✦	✓	✓
Working with SMCs, PTAs and other stakeholders	10	✦	✓	✓	✓	✓		✓	✦	✓	✦	✓	✓	✓	✓
Total	13	✓	✓	✓	✓	✓	✓	✓	✓	✓	✦	✓	✓	✓	✓

Key: ✓ = The intervention is at the core of the project's intervention strategy. ✦ = The intervention is being used, but is not a core activity.
Source: Adapted from the FM's intervention mapping, accounting for updates provided in January 2017.

Midline evaluation findings

The EM household survey, school visit survey and school-based assessment (SBA) data include a limited number of variables that can help us assess school governance. The household survey asks families whether they are involved in any school committees. The SBA asks school teachers whether or not they think there is enough contact between parents and the school. We added more specific questions about school development planning at midline, but have no baseline-to-midline comparison for these new variables. We will be able to analyse them at endline.

As shown in Table 30, we did not measure any significant effect on families' involvement in school management committees, or the extent to which parents and schools interact according to teachers. At baseline, 82 per cent of households said that there was no school committee at the girl's school or that they were not involved. At midline, the share was roughly the same (80%) in both the treatment and control group (82% / 80%, 85% / 80%, 3 %-p). At the same time, school committees typically involve only a relatively small number of parents from a given school, and parents may well be engaged in their children's education without being part of a committee. At midline, we

added new questions to the School Visit Survey to gauge changes in school governance (e.g. whether the school has a development plan in place) more directly at endline.

The share of teachers stating that there was not enough interaction between the school and parents remained virtually unchanged in the treatment group but increased by 10 percentage-points in the control group. However, the difference-in-difference is not statistically significant (36% / 38%, 37% / 47%, -8 %-p).

Only one project, ACTED (Afghanistan) shows a significant, albeit very small effect on families' likelihood to be involved in school governance and management. The share of families not involved in school governance increased slightly from 70 per cent to 72 per cent in the treatment group, but it increased even more in the control group (70% / 72%, 69% / 73%, -2* %-p). Interestingly, analysis of project data shows that the proportion of households reporting that activities to improve school management have been carried out since baseline is highest compared to other SCW project areas (73% in ACTED treatment areas, while it is less than one third of households in other project areas). This may be due to the involvement of community and religious leaders in schools.

Table 30: Difference-in-Difference in barriers relating to school management and governance

Difference-in-difference by project EM 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 Z-T
HHS															
No school committee or household members not involved	3%	2%	4%	-2%	4%	3%	0%	4%	0%	6%	4%	5%	11%	0%	
SBA															
Not enough contact between parents and school	-8%				4%	-32%	-20%	6%	11%						

Budget expenditure and effectiveness

As shown in Table 31, projects spend on average 12 per cent on school governance and management. However, this average conceals considerable variation across projects, with Save the Children (Mozambique) investing 37 per cent of their budget on school governance (see Box 11), and four projects not reporting any budget line for this intervention category. The latter included ACTED who are the only project showing a statistically significant (albeit small) effect.

Table 31: Percentage of the project budget spent on school governance and management

SCHOOL MANAGEMENT & GOVERNANCE	Avg. spent	BRAC Afg	AKF Afg	Acted Afg	IRC DRC	STC Eth	ChHp Eth	WUSC Ken	CfBT Ken	STC Moz	Plan Sie	RI Som	CARE Som	WV Zim	Camfd Z-T
Community and private schooling provision	0%	*	*						4%		N/A				
Technology for school management	1%										N/A				6%
Working with local or national education authorities	7%		3%						2%	37%	N/A	22%	13%	*	
Working with SMCs, PTAs and other stakeholders	4%		15%		4%				3%		N/A	2%	14%	2%	11%
Total investment	12%	0%	19%	0%	4%	0%	0%	0%	8%	37%	N/A	24%	27%	2%	16%

Note: *Note that for BRAC and AKF, establishing community-based schools is at the core of their theory of change. Costs relating to this overarching activity are subsumed by other budget lines such as 'teachers recruited' or 'school and classroom building/improvement' (see relevant tables above), which explains why they do not appear in this table. World Vision has worked with local and national education authorities to deliver different aspects of their project even though this is not reflected in the available budgetary data.

Source: Budgetary information and assigned categories were extracted from the FM's Value for Money analysis, and adapted if needed.

Box 11: Investing in school governance the example of Save the Children (Mozambique)

As shown in Table 31, Save the Children (Mozambique) spent the most on school governance and management. Save the Children (STC) works towards two outputs that explicitly address issues with school governance. The first is "safer and more gender-sensitive learning environments and improved school management" and the second is "increased commitment from government stakeholders to sustain improvements to girls' education". As such, they are one of the only SCW projects to build sustainability explicitly into their theory of change.

To promote safer, more gender friendly schools and improved school management, STC establishes "safe school committees" and supports them in developing action plans that outline how the school's safety could be improved. They also work with schools towards a set of "guiding principles for creating a quality learning environment" where the children's well-being is protected, their active learning encouraged, and parents and the community are involved in improving education. As many schools were faring poorly with regards to safeguarding children's wellbeing and engaging parents, STC have put in place a Child Protection Plan to address this issue. The project is also piloting a complaint response mechanism in intervention schools that allows girls to raise any incidences of inappropriate treatment or harassment that they may have experienced or witnessed.

To increase commitment from government stakeholders and sustain improvements to girls' education, STC have involved government officials in meetings to discuss issues such as access to education for vulnerable children, gender, quality of learning environment, and literacy. They have also trained government officials (e.g. focal points, school committees, school directors) so that they can build institutional capacity around gender, literacy and numeracy, safeguarding and support to vulnerable children.

The EM's qualitative evidence suggests that parents and school committees have been mobilised, collect money for school improvements, and are generally more engaged with school governance. This was described by one caregiver as follows:

> *Yes, we made a contribution to build new classrooms. We did contribute 500 kwachas (Malawi currency[61]); there was another phase when we contributed 1000 kwachas. This was a community initiative in coordination with the school directorate. This was successful because we added more classrooms*.
>
> (IDI with a caregiver, Nkhame, STC Mozambique).

One community leader said that the community now holds the local school to account for their management of financial resources: "*we question the substance of what they have done with the money*" (IDI with a Community Leader, Nkhame, STC, Mozambique).

3.4.5 Community attitudes and aspirations

Two thirds of projects assumed in their theories of change that negative parental and community attitudes towards girls' education led to girls being less likely to enrol and remain in school than boys.

Summary of baseline findings

The EM's qualitative baseline research found that, in some contexts, traditional perceptions about the role of girls and considerations around the economic return of educating a girl as opposed to marrying her off for a dowry led families to prioritise marriage over education once girls reached secondary school age. The EM's quantitative analysis showed a significant relationship between negative attitudes and lower education outcomes. But a key baseline finding was that the prevalence of negative attitudes was generally lower than many projects had anticipated, which has been further confirmed by the EM thematic research.

Community based interventions to improve community attitudes and aspirations

Projects have undertaken a variety of activities to raise awareness of the value of education in communities, mobilised parents to (re-) enrol their girls in school, and engaged communities in improving school governance and support structures for girls (see

Table 32). All SCW projects except for Plan (Sierra Leone) had some kind of community-based intervention at the core of their theory of change. Eight projects worked with parent or women groups, four projects organised community meetings to discuss girls' education issues, and three projects ran media campaigns or worked with faith groups or traditional leaders to advocate for the value of educating girls. A few other projects conducted similar activities, albeit not as part of their core theory of change.

[61] The community is on the border with Malawi, which may explain why the community uses Malawian currency.

Table 32: Number of SCW projects carrying out interventions to improve community attitudes

COMMUNITY BASED	Total core	BRAC Afg	AKF Afg	Acted Afg	IRC DRC	STC Eth	ChHp Eth	WUSC Ken	CfBT Ken	STC Moz	Plan Sie	RI Som	CARE Som	WV Zim	Camfd Z-T
Community meetings/ gatherings	4	✓	✦	✦	✦	✓	✓		✓			✦	✦	✦	
Household-level visits and support	1		✦	✦	✓					✦		✦	✦	✦	
Media campaigning (radio, TV, ads)	3		✦	✓	✦			✓		✦	✦	✓	✦		
Parents' and women's groups	8	✓	✓	✓						✓		✓	✓	✓	✓
Working with faith groups traditional leaders	3		✦									✓	✓	✓	
Working with men and boys	1		✦					✓	✦			✦	✦	✦	
Total	13	✓	✓	✓	✓	✓	✓	✓	✓	✓	✦	✓	✓	✓	✓

Key: ✓ = The intervention is at the core of the project's intervention strategy. ✦ = The intervention is being used, but is not a core activity.
Source: Adapted from the FM's intervention mapping.

Midline evaluation findings

The attitudes of primary caregivers toward girls' education are relatively positive in both treatment and control communities, with only a small percentage of households believing that education does not help people make their lives better and a similarly small percentage of households that do not want their girls to progress beyond primary school at midline.

Community and parental beliefs tend to be much more negative in BRAC (Afghanistan) and Save the Children (Ethiopia) areas, where more than half of the households do not usually send girls to school. In comparison to other projects a larger percentage of households surveyed by Save the Children do not want their girl to go beyond primary level, believe that girls learn less than boys at school, and think that it would be better for a girl to be married or working than in school.

Table 33: Midline project data on barriers related to community attitudes and aspirations

Midline barriers by project Project data		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
HHS															
PCG would spend more money on education for boys than girls	T	30%				5%	16%	20%	4%			3%	3%	3%	
	C					2%	12%		3%					3%	
PCG doesn't want girl to go beyond primary level	T	21%				56%	8%		4%			6%	2%	1%	
	C						8%		3%					1%	
PCG believes girls learn less than boys at school	T	10%			10%	24%	10%	12%	10%			3%	4%	4%	
	C				13%	22%	12%		9%					3%	
PCG believes it would be better for girl to be married or working than in school	T	21%				44%	9%	10%	4%			6%	3%	11%	
	C					38%	10%		4%					12%	
No school committee or household members not involved	T	70%					85%	86%	82%				77%	92%	
	C						86%		82%					92%	
Most/some people in the village don't send girls to school	T	65%				56%	4%	38%	27%			12%	15%	22%	
	C					50%	5%		24%					21%	

Key: Darker shades of red indicate higher prevalence of the related barrier.

Table 34 examines whether parental aspirations, community support for educating girls, and the participation of parents and communities in school governance have improved as a result of the above project interventions.

We find no significant improvement in parental attitudes and aspirations at the window level. The share of households that state that they spend more on the education of their boys than that of their girls has remained stable (13% / 12%, 11% / 10%, 0 %-p). The share of caregivers who would not like their girl to complete any higher education than primary level has increased marginally in both the treatment and control group (17% / 19%, 17 /

18%, 1 %-p). The share of caregivers stating that girls learn less than boys in school has similarly remained stable in both groups (12% / 11%, 11% / 11 %, -1 %-p).

While these barriers have not changed, **low parental aspirations and negative attitudes** are relatively rare in SCW project areas. Between 80 per cent and 90 per cent of caregivers provided supportive answers to these questions, suggesting rather positive attitudes towards educating a girl. During the EM's qualitative thematic research in three SCW project areas, researchers also found the same generally positive attitude to girls' education[62]. This may reflect government and NGO awareness-raising activities designed to increase demand for girls' education, or it may be an instance of respondents providing what they perceive are normative answers to questions. Nevertheless, it seems that attitudes are overall more positive than expected, but that engrained negative attitudes persist in some families and communities.

Table 34: Difference-in-Difference in barriers relating to community attitudes and aspirations

Difference-in-difference by project EM 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 Z-T
HHS															
PCG would spend more money on education for boys than girls	0%	-2%	2%	0%	-2%	-2%	0%	1%	0%	-1%	4%	-3%	-2%	1%	
PCG doesn't want girl to go beyond primary level	1%	5%	-3%	-1%	-3%	6%	2%	13%	1%	-5%	1%	-3%	-11%	0%	
PCG believes girls learn less than boys at school	-1%	-5%	2%	-13%	-8%	4%	-7%	-3%	-3%	6%	15%	3%	2%	4%	
Most/Some people in the village don't send girls to school	-4%	-1%	-19%	-6%	-16%	10%	-5%	-5%	4%	6%	-1%	-9%	3%	-9%	
Not enough support in community for girls to succeed in school	-5%	-1%	-10%	-5%	-16%	-5%	-23%	7%	-1%	-9%	1%	-2%	30%	-15%	

Key: The difference-in-difference indicator is equal to: DID = (Midline Treatment – Midline Control) – (Baseline Treatment – Baseline Control). The way variables are coded implies that *negative* DID demonstrate a *decrease of barriers to education*, hence a "positive" effect of the SCW.
Light green = barrier decreased significantly at the 10% level. Light orange = barrier increased significantly at the 10% level.
Darker green = barrier decreased significantly at the 5% level. Darker orange = barrier increased significantly at the 5% level.
Grey = no figure or sample size smaller than 50 individuals per project area per wave.

IRC (DRC) is the only project to show a significant positive effect on parental attitudes and aspirations. In the treatment group, the share of caregivers reporting that girls tend to learn less in school than boys remained stable in the treatment group whereas it increased substantially in the control group (5% / 5%, 3% / 11%, -8** %-p). It is unclear, whether this reflects a real worsening in attitudes in the control group, or whether this difference may have been caused by high attrition rates in IRC areas. Qualitative analysis suggest that caregivers seem to have similar perceptions of girls' and boys' ability. Respondents generally mentioned that girls are hard-working at school, but reported problems relating to early marriage and pregnancy, either resulting from parents seeking dowry or girls' own choices.

Box 12: The lack of secondary schools in WUSC refugee camps

In WUSC treatment areas, the share of caregivers not wanting their girl to go beyond primary school has soared from 7% at baseline to 24% at midline, while increasing much less (by 5 percentage points) in the control group. This might reflect a shortage of secondary schools for primary school leavers in the camps:

> *"Only about 50 per cent of the eligible 2,753 (567 female) candidates will be able to secure places in secondary schools unless strategies to increase secondary school places are agreed. This transition rate will be a significant drop from the 92 per cent transition level achieved in 2015 and will have devastating effect on student morale and performance in the future*." (UNHCR, 2015 Kenya Certificate of Primary Education (KCPE) Result Analysis, Kakuma Camps, WUSC's midline evaluation report)

In WUSC's midline evaluation report, the external evaluator equally stressed the uncertain future of girls who had done particularly well in primary school, but are now unable find a place at secondary school, or pay the associated costs:

> *"Parents who claim to be in favour of an education for their daughter may nevertheless arrange a marriage for her if the offer is right. Girls may imagine a more immediately brighter future in the*

[62] Evaluation Manager Girls' Education Challenge Fund (2016). Narrow Windows and Revolving Doors. GEC Thematic Research Report, January.

> *arms of a man with a bit of money to spend on her, than waiting in hopes of accessing a faraway classroom in a year or two."* (WUSC midline evaluation report, p. 62)
>
> This example underscores the importance of building sustainability into the project's theory of change from the start. Raising expectations and aspirations that cannot actually be met can cause harm to girls and families who have difficult decisions to make as they cope with very limited resources.

We find a significant effect on **community attitudes and support for girls' education** at the window level. At baseline, 17 per cent of respondents in the treatment group said that "some or most people do *not* usually send girls to school" and the share was similar in the control group. Even though this share had increased at midline, it had increased significantly less in the treatment group than in the control group (17% / 22%, 16% / 25%, -4** %-p). In addition, the share of caregivers stating that there was not enough support in the community for girls to succeed in school dropped at midline, while remaining stable in the control group (47% / 42%, 48% / 48%, -5* %-p).

This positive effect on community attitudes and support is driven by a number of projects. AKF (Afghanistan), IRC (DRC) and ChildHope (Ethiopia) show significant and (relatively) large effects for both variables. In AKF treatment areas, the share of caregivers reporting that it is unusual for parents to send their girls to school decreased from 37 per cent at baseline to 21 per cent at midline but increased in the control group (37% / 21%, 41% / 45%, -19** %-p). In IRC treatment areas, the share of caregivers reporting that there was not enough community support for girls' education fell at midline while increasing in the control group (40% / 32%, 37% / 45%, -16** %-p). Likewise, in ChildHope treatment areas, the share of caregivers reporting a lack of community support was almost halved in the treatment group, whereas it increased in the control group (28% / 16%, 2% / 32%, -23** %-p).

As shown in

Table 32 these projects used different strategies to mobilise communities and raise awareness of the value of girls' education. AKF (Afghanistan) mobilised communities to take on new schools after the end of the project, and used radio broadcasting and text messaging to advocate for girls' education in communities and at the national level. They also mobilised parents to participate in open school days, reading events, library groups and social audits. IRC delivered community information campaigns with messages promoting on-time enrolment, championing the importance of education for girls and boys, and combating socio-cultural barriers to girls' education. Our qualitative analysis found that community members have observed these activities, particularly the involvement of community chiefs, flyers' distribution, and diffusion of radio programmes. As this caregiver explained:

> *"We spoke on the radio, there were commercials and plays. That was a way of making it known, raising people's awareness on the importance of studies. I also listened to a recitation "The Two Roads", about a boy who at the end of a road finds two roads… He began to rub his hands, the sign was there but he couldn't read it. Why, because he never cared about studies. And we did it to raise parents' and children's' awareness on the importance of studies."* (Caregiver, Kasai-Orental, IRC, DRC)

ChildHope (Ethiopia) conducted community conversations with 4,200 community members to engage them in supporting the education of marginalised girls. Our qualitative analysis found that income grants and material support have also helped in changing attitudes and generating support to education in ChildHope project areas, but the most impactful activity seems to have been the involvement of religious leaders to raise awareness on a broad range of topics. Muslim and Christian leaders are perceived as powerful and respected voices within the community and appear to have been efficiently relaying messages on girls' education, opposing early marriage and abductions, and migration to the Middle East. School officials report going to churches on weekends and being allowed to take the stage and ask parents to send their children to school.

> *"Here when religious leaders ordered their followers to send your children to school at the age of seven, followers considered it as God's word. Therefore, everybody strictly applies what they have been said."* (IDI, Community Leader, Amhara, Childhope, Ethiopia)

Budget expenditure and effectiveness

As shown in Table 35 SCW projects are spending 11 per cent on average on community-based activities to improve attitudes and aspirations. World Vision (Zimbabwe) spent as much as 32 per cent of their budget in this area, whereas the budgetary data does not record any spending on community-based activities for Save the

Children (Ethiopia) and Camfed (Tanzania/Zimbabwe). This may be due to the specific categorisation of activities in the FM's budgetary mapping or due to missing budgetary data, as Camfed has invested in training for mother support groups (including financial management), and has provided start-up funds for some mother support groups to undertake income-generating activities.

AKF (Afghanistan) and ChildHope (Ethiopia), alongside IRC (DRC), achieved significant improvements in community support and attitudes. These three projects invested comparatively little in interventions focused on changing community based attitudes (5%, 8% and 6%, respectively). Interestingly, they all spent almost all of this budget in community awareness activities like meetings and gatherings, which suggest that these interventions could be particularly effective at improving target communities' attitudes and support to girls' education. It is worth nothing that the AKF project midline evaluation report stressed that a lack of a clear hand-over strategy for the project may have led to disillusionment in the communities and led to resentment and negative attitudes. However, as discussed above, this is not being picked up by our EM data which shows positive improvement in attitudes in AKF treatment areas.

We observe no significant improvement for CfBT (Kenya) and World Vision (Zimbabwe) project areas, even though these projects invested the most in community-based interventions. World Vision treatment areas do show a difference-in-difference improvement, compared with the control group, but the effect is not statistically significant. ChildHope's and CfBT's midline evaluation reports both highlighted that community conversations were slow-starting and less effective than expected, because communities had both deeply engrained views that were slow to change, as well as other priorities that often took precedence over discussing issues related to girls' education.

Table 35: Percentage of the project budget spent on community based activities

COMMUNITY BASED	Avg. spent	BRAC	AKF	Acted	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
Adult literacy	0%				0.4%						N/A		4%		
Community meetings/ gatherings	5%	1%		5%	5%		7%		26%	4%	N/A		6%	6%	
Household-level visits and support	1%		2%		0.3%		1%	6%	3%	1%	N/A				
Media (radio, TV, advertising)	1%		3%							1%	N/A	3%			
Parents' and women's groups	1%									3%	N/A	1%		4%	
Working with faith groups / traditional leaders	3%	2%									N/A		4%	23%	
Working with men and boys	1%							6%			N/A				
Total investment	11%	3%	5%	5%	6%	0%	8%	12%	30%	9%	N/A	4%	15%	32%	0%

Source: Budgetary information and assigned categories were extracted from the FM's Value for Money analysis, and adapted if needed.

3.4.6 Violence

Twelve SCW projects assumed in their theories of change that issues around safety and violence function as barriers to girls' education. This included general insecurity, conflict, (domestic) violence, sexual harassment, assault or violence, and corporal punishment.

Summary of EM baseline findings

Evidence from the qualitative research at baseline suggested that issues around **insecurity and violence** were a common concern among families in SCW project areas and affect girls' attendance and learning. IDI respondents mentioned different forms of violence (e.g. political conflict, active inter-tribal conflict, domestic violence, or sexual harassment) and different effects on girls' education, ranging from the physical destruction of schools to psychological trauma or aggressive behaviour in school.

Due to the highly sensitive nature of issues around violence, and sexual violence in particular, we have very limited quantitative evidence about this barrier. The household survey suggests that a significant proportion of girls (24%) have unsafe journeys to school and that this affects their attendance. However, the baseline research had limited capacity to explore the prevalence and effects of violence at home, on the way to and from school, at school and in the communities.

Interventions that address violence against children

In education projects, the scope to tackle the root causes of conflict and violence or prevent their occurrence is usually limited. As shown in Table 36 seven SCW projects have made addressing violence against children a key part of their theory of change and use different strategies to fight it. Five projects are raising community awareness around issues related to violence, three projects have developed child protection policies in schools, and one project has explicitly strengthened referral pathways in school. World Vision (Zimbabwe) and Save the Children (Mozambique) also explicitly work to reduce harmful practice such as child marriage or female genital mutilation (FGM). Other projects also report that they are delivering interventions that address violence, but they are not core to their theory of change.

Table 36: Number of SCW projects that are explicitly tackling violence against children

VIOLENCE-RELATED	Total core	BRAC Afg	AKF Afg	Acted Afg	IRC DRC	STC Eth	ChHp Eth	WUSC Ken	CfBT Ken	STC Moz	Plan Sie	RI Som	CARE Som	WV Zim	Camf Z-T
Addressing abuse from adults in charge	0		✦											✦	
Addressing harmful practice (e.g. child marriage, FGM)	2									✓		✦	✦	✓	
Addressing corporal punishment	1		✦				✓					✦	✦		
Addressing violence between children	0	✦													
Community awareness	5		✓	✦		✓				✦		✓	✓	✓	
Development of child protection policies in schools	3				✦	✦	✓			✓		✦		✓	✦
Strengthening referral paths	1		✦			✦	✦			✦		✦		✓	
Total	7	✦	✓	✦	✦	✓	✓			✓		✓	✓	✓	✦

Key: ✓ = The intervention is at the core of the project's intervention strategy. ✦ = The intervention is being used, but is not a core activity.
Source: Adapted from the FM's intervention mapping.

Midline evaluation findings

The EM midline data does not show any effect on barriers relating to violence at the window level (see

Table 37). The share of caregivers reporting that it was fairly or very unsafe to travel to school in the area remained at 14 per cent, and at similar levels in the control group. The share of caregivers reporting that the girl had a dangerous experience while traveling around the area increased slightly from 2 per cent to 5 per cent in the treatment group (2% / 5%, 4% / 6%, 1 %-p), potentially reflecting that girls are at a greater risk of harassment when they reach adolescence or that there are fewer secondary schools than primary schools, so girls' distance to school is greater. The share of respondents reporting that there has been a violent incident at the girl's school in the past year increased from 3 per cent to 7 per cent in the treatment group, and to a similar extent in the control group (3% / 7%, 3% / 6%, 0 %-p).

Care (Somalia) is the only project where significantly fewer families reported dangerous experiences while travelling around in the treatment group than in the control group. Our qualitative research found evidence of reported violence in the communities in the project areas. In the cases where it was mentioned, it was largely due to external factors. Insecurity, armed conflicts and clashes between clans were reported, such as in the area of Galkaio, and have had a dramatic impact on education with respondents reporting schools being destroyed. However, a high number of respondents reported an improvement compared with past years, describing safe spaces where they send their children to school without fear. Besides, our qualitative evidence found that teacher sensitivity trainings run by the project have been useful as teachers in some areas have been reported using alternative punishments such as making students do push-ups, and watering the plants and trees at school.

The share of teachers reporting issues with safety and security at school has increased across all SBA treatment areas except Save the Children (Ethiopia). This is likely due to external factors that projects could not directly impact on. Teachers' endorsement of corporal punishment at school has also slightly increased overall in comparison with control areas, in the same proportions and in the same project areas where issues with safety have been reported.

Table 37: Difference-in-Difference in barriers relating to violence

Difference-in-difference by project EM 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 Z-T
HHS															
Fairly or very unsafe travel to schools in the area	1%	0%	5%	-1%	1%	2%	-2%	0%	2%	1%	-1%	1%	0%	2%	
Girl has had dangerous experience travelling around	1%	-3%	3%	-2%	2%	1%	0%	-3%	-1%	-1%	2%	8%	-11%	5%	
Violence at girl's school in the past year	0%	-1%	-6%	-3%		2%	2%	-4%	2%		-2%	7%	-2%	6%	
SBA															
Issues with safety at school	18%					-16%	35%	38%	20%						
Teacher feels corporal punishment is appropriate	12%				1%	-4%	22%	20%	17%						

Key: The difference-in-difference indicator is equal to: DID = (Midline Treatment – Midline Control) – (Baseline Treatment – Baseline Control). The way variables are coded implies that *negative* DID demonstrate a *decrease of barriers to education*, hence a "positive" effect of the SCW. Light green = barrier decreased significantly at the 10% level. Light orange = barrier increased significantly at the 10% level. Darker green = barrier decreased significantly at the 5% level. Darker orange = barrier increased significantly at the 5% level. Grey = no figure or sample size smaller than 50 individuals per project area per wave.

Budget expenditure and effectiveness

No SCW projects explicitly stated what proportion of their budget they had invested in addressing violence against children. This is most likely due to only a small share being spent specifically on these types of activities, or that they were built into other activities. Interventions aimed at tackling violence in school are often integrated with community awareness activities, or with teachers' training interventions as in the case of CARE (Somalia) where teachers where trained on child protection.

3.4.7 Personal factors

Twelve SCW projects assumed that personal factors acted as barriers to girls' education. They focused on different issues such as early pregnancy, early marriage, disability, girls' self-esteem, aspirations and motivation, the level of education in the family, and migration and mobility.

Summary of EM baseline findings

Early motherhood and pregnancy along with **early or forced marriage** were often mentioned as barriers to education in IDIs. In particular, respondents described the stigmatisation of young mothers, which often leads to them leaving school. Respondents also mentioned girls leaving school deliberately to start a family based on considerations about viable economic pathways rather than viewing it as a failure or drop-out.

In the IDIs, **disability** was not generally viewed as a barrier to children's participation in education. The quantitative evidence showed that some forms of disability were marginally associated with lower attendance and learning, but it is likely that disability was under-reported in both our quantitative and qualitative data.

Low parental education was found to be associated with girls' educational outcomes. This is in line with international literature about education and the intergenerational transmission of educational disadvantage, but low parental education was not specifically cited by projects as a barrier to education.

No quantitative evidence was available to link **mobility and migration** with poor education outcomes controlling for other factors, although reading levels in some pastoralist contexts were particularly low. However, IDI respondents mentioned migration and resettlement as causes of poor outcomes due to disruption of enrolment and attendance.

Interventions addressing personal factors

Many of the personal factors mentioned above are structural and / or difficult to change. For example, disability, low parental education or a history of migration are factors that have underlying causes that projects will not be able to address directly. At best, project interventions are most likely to remedy the effects that these factors have on the chances of marginalised girls being in school and learning effectively.

Four projects are delivering activities to empower girls and improve their self-esteem as a core part of their intervention strategies (see Table 38). They promote girls' voice and participation, provide mentoring, role models or safe spaces. However, the effects of such activities on girls' self-esteem are difficult to measure through the EM's household survey which is not designed for this purpose.

Table 38: Number of SCW projects carrying out activities to improve self esteem or empower girls

EMPOWERMENT & SELF-ESTEEM	Total core	BRAC	AKF	Acted	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
Activities that promote girls' voice and participation	3		✓	✦						✓			✦	✓	
Mentoring	3	✓	✦				✓								✓
Role models (older girls, female teachers, parents)	1		✦	✦								✦			✓
Safe spaces	2		✓	✦						✓			✦	✦	
Total	4	✓	✓	✦						✓			✦	✦	✓

Key: ✓ = The intervention is at the core of the project's intervention strategy. ✦ = The intervention is being used, but is not a core activity.
Source: Adapted from the FM's intervention mapping.

Six projects are addressing some of the above-mentioned personal factors that can exacerbate the marginalisation of girls from education (see Table 39). Four projects are working with remote or nomadic communities, including AKF and ACTED (Afghanistan), Save the Children (Ethiopia) and WUSC (Kenya). Plan (Sierra Leone) addresses some of the barriers related to being disabled. Using EM quantitative data, it is difficult to assess the effectiveness of these interventions as they do not directly relate to the change of any specific variable from our surveys. Besides, many projects have addressed personal factors through a range of different activities, such as early marriage and pregnancy through community-awareness raising interventions.

Table 39: Number of SCW projects carrying out interventions with particularly marginalised groups of girls

MARGINALISATION-RELATED	Total core	BRAC	AKF	Acted	IRC	STC	ChHp	WUSC	CfBT	STC	Plan	RI	CARE	WV	Camfd
		Afg	Afg	Afg	DRC	Eth	Eth	Ken	Ken	Moz	Sie	Som	Som	Zim	Z-T
Interventions in remote or nomadic locations	4	✦	✓	✓		✓		✓	✦				✦		
Interventions addressing cultural/ linguistic exclusion	1					✓			✦						
Interventions addressing disability	1	✦									✓				
Interventions with other marginalised groups	3						✓							✓	✓
Total	6	✦				✓	✓	✓	✦		✓		✦	✓	✓

Key: ✓ = The intervention is at the core of the project's intervention strategy. ✦ = The intervention is being used, but is not a core activity.
Source: Adapted from the FM's intervention mapping.

All projects provide opportunities for extra-curricular and non-formal learning as part of their intervention strategy (see Table 40). Projects teach girls numeracy and literacy, life skills, or vocational skills, and provide mentoring and guidance. In relation to personal factors, these interventions cannot be linked directly to a specific set of barrier variables from our surveys. They represent a remedy to insufficient teaching in school, often targeting marginalised groups, such as out-of-school girls or girls who perform particularly poorly in school, and aim to enhance learning skills directly.

Table 40: Number of SCW projects that offer extra-curricular or non-formal education opportunities

EXTRA-CURRICULAR & NON-FORMAL EDUCATION	Total core	BRAC	AKF	Acted	IRC	STC	ChHp	WUSC	CfBT	STC	Plan	RI	CARE	WV	Camfd
		Afg	Afg	Afg	DRC	Eth	Eth	Ken	Ken	Moz	Sie	Som	Som	Zim	Z-T
Life skills (including sexual and reproductive health)	6					✓	✓		✓	✓	✦	✦		✓	✓
Mentoring (peer support, learner guides)	4	✓					✓	✓		✦			✦	✦	✓

Mixed sex/ additional boys' clubs	4		✦			✓	✦	✓	✓			✦	✦	✦	✓
Non-formal / alternative education	3		✓	✓	✓										
Tutoring (homework clubs, reading/ literacy clubs)	8		✦	✓	✓		✓	✓		✓	✓	✓	✦		✓
Vocational training & economic empowerment	2			✓							✓				
Total	13	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✦	✓	✓

Key: ✓ = The intervention is at the core of the project's intervention strategy. ✦ = The intervention is being used, but is not a core activity.
Source: Adapted from the FM's intervention mapping.

Midline evaluation findings

Many of the personal or family factors mentioned above are structural and unlikely to change through project interventions. However, many projects do try to improve girls' aspirations, self-confidence and skills through extracurricular activities and non-formal education. As shown in Table 41, we do not find any significant change in girls reporting that they enjoy school, or in their attitudes towards going to school. It is worth noting, however, that only very few girls (2% -3%) at baseline stated that going to school was not important or not good in the first place so it was very difficult to measure a significant improvement.

In WUSC (Kenya) project areas, the share of girls who said that they did not enjoy school or only sometimes rose at midline, whereas it decreased in control areas (15% / 21%, 16% / 8%, +14** %-p). The situation worsened significantly in the treatment group, compared with the control group. The reason is not clear from our qualitative research but it may be due to increased expectations from the school environment rather than to an absolute decrease in girls' aspirations.

IRC (DRC) and Save the Children (Ethiopia) show a large although not significant improvement in the share of girls (in grades 4 and 6) who reported not feeling confident doing maths, which is reflected in the significant decrease across the five project areas covered by the SBA (12% / 1%, 7% / 3%, -7* %-p). This suggests that projects' activities have made a tangible contribution to strengthening girls' self-confidence in doing maths.

Table 41: Difference-in-Difference in barriers relating to girls' aspirations and self-confidence

Difference-in-difference by project EM data	All	BRAC Afg	AKF Afg	ACTED Afg	IRC DRC	STC Eth	ChHp Eth	WUSC Ken	CfBT Ken	STC Moz	Plan Sie	RI Som	CARE Som	WV Zim	Camf Z-T
HHS															
Girls doesn't like school or only sometimes	-1%	7%	-7%	-12%	2%	-6%	-5%	14%	1%	-2%	-8%	-2%	-8%	2%	0%
Girl doesn't think it is good for children to go to school	-2%	0%		-4%	0%	3%	-1%		-1%	-4%	-3%		-5%	-1%	0%
Girl doesn't think school is important for her future	1%	1%	3%	0%	1%	-5%	0%	2%	3%	-3%	0%	-1%	7%	-2%	0%
SBA															
Girl does not feel confident reading and writing	-2%				3%	-13%	0%	0%	0%						
Girl does not feel confident about doing mathematics	-7%				-12%	-18%	0%	3%	-5%						

Key: The difference-in-difference indicator is equal to: DID = (Midline Treatment – Midline Control) – (Baseline Treatment – Baseline Control). The way variables are coded implies that *negative* DID demonstrate a *decrease of barriers to education*, hence a "positive" effect of the SCW.
Light green = barrier decreased significantly at the 10% level. Light orange = barrier increased significantly at the 10% level.
Darker green = barrier decreased significantly at the 5% level. Darker orange = barrier increased significantly at the 5% level.
Grey = no figure or sample size smaller than 50 individuals per project area per wave.

The variables measuring girls' aspirations that we present above are more suited to girls of an early age. We have included more nuanced questions measuring aspirations, motivation and self-confidence in the midline survey and will be able to look at changes in these more specific variables in our endline evaluation.

Budget expenditure and effectiveness

In this section we are considering only the budget spent by projects on extra-curricular activities, in the absence of detailed budget figures for empowerment / self-esteem and marginalisation-related activities.

On average across the window, projects spent 14 per cent of their budget on extracurricular activities and alternative education (see Table 42). ACTED (Afghanistan) stands out for investing half of their overall budget in this category, building youth development centres that offer literacy courses and vocational training (see Box 13). Other projects, like Care (Somalia), WUSC (Kenya), RI (Somalia), or Camfed (Tanzania and Zimbabwe) invested only a small fraction of their budget on alternative and non-formal education.

Table 42: Percentage of the project budget spent on extracurricular activities and non-formal education

EXTRA-CURRICULAR & NON-FORMAL EDUCATION	Avg. spent	BRAC	AKF	ACTED	IRC	STC	ChHp	WUSC	CfBT	STC	Plan	RI	CARE	WV	Camfd
		Afg	Afg	Afg	DRC	Eth	Eth	Ken	Ken	Moz	Sie	Som	Som	Zim	Z-T
Life skills (including sexual and reproductive health)	3%								13%		N/A			18%	
Mentoring (peer support, learner guides)	2%	18%					1%	2%			N/A	2%			2%
Mixed sex/ additional boys' clubs	0%					2%	0.4%		4%		N/A				
Non-formal / alternative education	4%		3%	37%	11%						N/A				
Tutoring (homework clubs, reading/ literacy clubs)	4%		6%		5%	4%	14%			19%	N/A				
Vocational training & economic empowerment	2%		8%	13%							N/A				
Total investment	14%	18%	17%	49%	16%	6%	16%	2%	16%	19%	N/A	2%	0%	18%	2%

Source: Budgetary information and assigned categories were extracted from the FM's Value for Money analysis, and adapted if needed.

In ChildHope (Ethiopia) and CfBT (Kenya) project areas, we see significant improvements in girls' self-confidence with regards to reading and maths (respectively) both invested 16 per cent of their budget on extracurricular activities and non-formal education. ChildHope (Ethiopia) established girls' and "good brother" clubs, hired mentors and tutors, and offered tutorials for low-performing students. CfBT established girls' and boys' clubs and trained teachers, students and parents on behavioural change (gender equality, inclusion and adolescent health). CfBT also carried out activities on improving health and life skills. They ran health clubs for students, provided deworming tablets, vitamins & mineral supplements, and referred students for special needs support, care and treatment.

Many other projects spent equal amounts of their budgets on similar activities without showing any effects on girls' self-confidence as measured by the variables presented above. This underscores the difficulty of assessing what works and what does across a portfolio of 14 projects working in very different contexts. It is possible that some of the differences observed stem from some projects *delivering* the same activities more effectively. It is also possible that parts of the effects achieved by projects are not captured through our research instruments.

Box 13: ACTED's investment in alternative education

ACTED intervenes in Afghanistan's Faryab province, where girls are marginalised due to the remoteness of many communities, cultural restraints, gender inequalities and on-going conflict. Government provision of education is insufficient so that many girls have limited access to schools in their local area. ACTED focuses on providing alternative pathways for education. They have established 10 formal primary schools, and 7 Youth Development Centres for girls as well as village literacy courses. These three different modalities cater to the varying needs of girls with different ages and status of enrolment.

Youth Development Centres deliver literacy and vocational training courses and provide latrines, hand washing stations and basic safety equipment for emergencies. The project has recruited vocational trainers and constructed vocational workshops, distributed tool kits and materials for vocational training, hired nursery assistants, as well as community literacy teachers and materials.

While we cannot show any impact on girls' confidence and learning in our quantitative data, we saw some positive feedback in the qualitative data. Several respondents reported that girls had improved their skills and ability to read and write, as illustrated by the following quote:

> *Compared to the past years, now a lot of changes occurred in our children. A lot of changes occurred in their lessons and they have learned a lot of things. They can read signs well. They can spell their names easily; they can write wedding cards. Their talents have improved*.
>
> (IDI with a caregiver in Attaullah, ACTED Afghanistan)

The Youth Development Centres and Community Literacy Courses were praised by many respondents. Some reported that after attending the Community Literacy Classes girls have enrolled in the formal school. Others mentioned that providing literacy courses with the support of the community elders enabled married girls to attend lessons as the stigma was reduced. Several respondents stressed that the vocational training enabled girls to contribute to family income and afford the costs of formal schooling, as illustrated by the following quote:

> *My daughter was enrolled in embroidery and tailoring because she was older when she was recruited for the embroidery project. She earns money. She spends the money on her school expenses like school clothes*.
>
> (IDI with a caregiver in Khuja Qoshri Payan, ACTED Afghanistan)

Some respondents, however, stressed that the quality of teaching remained low, including at the Youth Development Centres, and that further training was needed.

Table 43 below provides a summary of the proportion of budget spent by SCW projects on each set of activities (blue rows). Below each budget row, we show the difference-in-different coefficients for the barriers that are the most relevant to those activities. Data has been extracted from tables previously shown in this section and aim at giving an overview of the way budget has been allocated by SCW projects as well as an indication of the effectiveness of their different activities on barriers to girls' education.

Table 43: Overview of the proportion of budget spent by each project across intervention categories and effect on related barriers

Intervention types	BRAC	AKF	Acted	IRC	STC	ChHpe	WUSC	CfBT	STC	Plan	RI	CARE	WV
	Afg	Afg	Afg	DRC	Eth	Eth	Ken	Ken	Moz	S-L	Som	Som	Zim
Economic	17%	8%	0%	54%	11%	42%	30%	18%	23%	N/A	25%	2%	39%
Difficult to afford to go to school				0%	-1%	-8%	-22%	4%	4%	8%	9%	4%	-6%
Teacher Training	53%	41%	0%	11%	1%	4%	11%	24%	9%	N/A	15%	10%	0%
Teaching is not satisfactory at girl's school	-10%	-35%	-15%	7%	-11%	-5%	9%	1%	-6%	4%	-16%	4%	1%
Teacher has no teaching certificate				3%	-6%	16%	-8%	11%					
School and classroom building/ improvement	6%	5%	45%	4%	61%	9%	14%	1%	2%	N/A	15%	36%	0%
Classrooms are not satisfactory	-5%	-29%	-25%	28%	-12%	-14%	-10%	-8%	-1%	-11%	-7%	17%	6%
School doesn't have electricity all day					7%	13%	28%	8%	4%		12%	-3%	-8%
School doesn't have a roof					-6%	3%	15%	7%	-5%		-12%	15%	1%
Textbooks & Learning materials	3%	4%	2%	6%	21%	5%	32%	1%	1%	N/A	11%	1%	0%
Textbooks are not satisfactory	8%	-25%	10%	22%	2%	-11%	-1%	-17%	1%	-1%	-27%	16%	2%
Children do not use textbooks in class				6%	-4%	0%	2%	1%	0%		2%	-1%	-1%
Toilets & WASH facilities	0%	0%	0%	0%	0%	17%	0%	2%	0%	N/A	4%	3%	8%
Toilets are not satisfactory at girl's school	4%	-38%	-25%	28%	4%	-2%	-12%	1%	-2%	7%	-13%	19%	6%
School doesn't have indoor toilets					-22%	11%	1%	3%	8%		-12%	-8%	-9%
School doesn't have access to water					11%	-1%	-2%	-1%	-26%		22%	23%	-8%
School governance	0%	19%	0%	4%	0%	0%	0%	8%	37%	N/A	24%	27%	2%
No school committee or household members not involved	2%	4%	-2%	4%	3%	0%	4%	0%	6%	4%	5%	11%	0%
Community attitudes	3%	5%	5%	6%	0%	8%	12%	30%	9%	N/A	4%	15%	32%
Most/some people in the village don't send girls to school	-1%	-19%	-6%	-16%	10%	-5%	-5%	4%	6%	-1%	-9%	3%	-9%
Not enough support in comm. for girls to succeed in school	-1%	-10%	-5%	-16%	-5%	-23%	7%	-1%	-9%	1%	-2%	30%	-15%
Thinking now, PCG doesn't want girl to go beyond primary	5%	-3%	-1%	-3%	6%	2%	13%	1%	-5%	1%	-3%	-11%	0%
Extra-curricular activities	18%	17%	49%	16%	6%	16%	2%	16%	19%	N/A	2%	0%	18%
(Literacy outcome) Girl's EGRA oral reading score	0	5	5	0	-1	1	1	14	-2	0	6	-7	-1
(Numeracy outcome) Girl's EGMA percent correct	1%	8%	2%	4%	-5%	2%	5%	4%	0%	4%	4%	-9%	-1%

Key: Blue rows show the proportion of budget spent by each project on each group of activities. Other rows show difference-in-difference estimators for related barriers. Colour coding is the same as in the section above: Green shows a significant decrease of the barrier at the 5% (darker) or 10% (lighter) level. Orange shows a significant increase of the barrier.
This **colour coding is inverted for the last two rows** showing difference-in-differences in learning outcomes: green shows a significant increase in score, orange shows a significant decrease.
Note that Camfed is not included due to the lack of a control group in the EM data.

Summary of key findings – What has worked, why and with what effects?

Despite the caveats described in our findings, we observe some interesting patterns regarding the effectiveness of different intervention types: projects which invested the most in economic interventions are also the ones that managed to substantially increase school affordability. The provision of bursaries and in-kind support, as well as loans in the case of World Vision (Zimbabwe), seems to have been an effective way of alleviating the burden of school-related expenses in some communities.

Similarly, two out of the four projects that spent most of their funding on teacher training show large improvements in caregivers' satisfaction with teaching: BRAC and AKF (both Afghanistan). Despite a slight deterioration in satisfaction in the treatment group, Relief International (Somalia) also record a highly positive difference-in-difference because the deterioration was much worse in the control group. CfBT (Kenya), spent 23 per cent of their budget on skills training, without showing a significant effect on perceptions of teacher quality in the EM data, even though their own data showed the highest reported increase in satisfaction with teaching out of any project.

It is important to note, however, that satisfaction with the quality of teaching does not necessarily translate into an actual improvement in teaching practice, or into improved learning outcomes for students. Parental perceptions are only one aspect to consider. Classroom observations and teacher assessments are required to draw a more comprehensive picture and the EM was not able to conduct these for all projects.

Across the SCW and on the basis of EM data, the most important barriers to girls' education identified at baseline in the target communities (poverty, school-based barriers) have not changed significantly as a result of GEC activities. Only a few barriers, mostly related to community attitudes and caregivers' satisfaction with school facilities, seem to have improved at the window level – often marginally.

However, **window-level variations hide important disparities across project areas**. Some projects seem to have performed much better than others. As shown in Table 43, the Aga Khan Foundation and ACTED (Afghanistan) have demonstrated a significant positive effect on learning outcomes at midline, along with CfBT (Kenya). These two projects have managed to significantly reduce barriers to girls' education across a variety of dimensions: poverty and livelihoods (Table 19) shows that they have had a positive effect on the time spent by girls on non-school work); teaching; learning materials; school facilities; school governance; and community attitudes. This general pattern of positive change has not been observed in CfBT (Kenya) areas, although the project has managed to largely improve the satisfaction of caregivers with textbooks, and substantially reduced the shortage of teachers in schools (see Table 23). CfBT (Kenya) provided teachers with special mentoring and training which may have had a positive impact on girls' learning that was not captured by our research instruments.

The majority of SCW projects have not had a significant positive effect on learning outcomes. Although they all spent substantial amounts of money on different types of intervention, they often were not able to lower barriers to girls' education as a result of their activities. **Most projects have had an effect on a few barriers, but either the size and scale of their impact was not large enough, or those barriers were not the most important (or impactful) constraints to girls' education**. In IRC (DRC) and ChildHope (Ethiopia) project areas, community attitudes towards girls' education have improved considerably since baseline (see Table 34). But this effect has not yet translated into any improvements in learning outcomes, possibly because other important factors such as economic or school-related barriers were not sufficiently addressed by midline.

Key lessons learned

- Projects that have had a positive effect on learning outcomes are those that managed **to effectively address a variety of direct and indirect barriers to education,** which include barriers that have been shown to have the most direct impact on girls' education (i.e. economic barriers and barriers related to poor teaching and school facilities). Examples of these projects are AKF and ACTED, which are both operating in Afghanistan. Projects which had a positive effect on just a few barriers have not managed to significantly improve learning outcomes.
- Evidence suggests that the most effective projects are **projects which were aware of the needs of their target populations from the beginning and who directly responded** to girls' and communities'

needs – for example, the three Afghan projects, and in particular AKF. The project helped set up community-based classes and schools in their target communities. Alongside this intervention, AKF carried out a range of activities to support teacher and school staff trainings and improve school infrastructure.

- The **size and significance of projects' effects on barriers often do not seem to be correlated to the proportion of budget spent on related activities**. This might suggest that the effect of a single type of intervention can vary a lot depending on the population targeted, the context in which the project operates, the other activities carried out alongside it, but also on the **quality of implementation**.
- Unfortunately, the lack of reliable evidence at the project level hindered our ability to evaluate *how well* projects managed and delivered their activities. Only very few projects managed to merge their baseline and midline household survey data, making it impossible to analyse changes over time in household, community, and school-level barriers. In projects' midline evaluation reports, the effectiveness sections were often very succinct and did not present a systematic assessment of changes in barriers, and effectiveness of interventions triangulating quantitative and qualitative data. We therefore had to rely mostly on EM data, which has not been designed to generate findings at the project level, and on value for money economy metrics from the FM. As a result, it has been extremely difficult to establish causal linkages between specific activities and barriers to education.

  The GEC has an overall Theory of Change and SCW projects have their own theories of change for the project as a whole. However, each project needs to make it clearer from the start **how and why specific types of interventions will address specific barriers** to girls' education. Intended causal pathways, as well as any mediating factors, such as those relating to the intervention context or delivery, need to be explained better in project design and evaluation reports, and assessed more systematically using all the available project evidence.

3.4.8 What were the unintended consequences of SCW interventions?

SCW projects anticipated a range of positive changes in barriers to girls' education in their theories of change and we have discussed many of these above. This section discussed the unintended consequences that SCW interventions have had on girls and their communities.

Resentment over the prioritisation of girls

The adverse effect most cited in project midline evaluation reports, and most frequently observed in the EM's qualitative interviews relates to resentment amongst boys (and their families) over the selective support provided to girls by SCW projects. Six project midline evaluation reports presented evidence that boys were feeling left behind and jealous of the stipends, textbooks, uniforms, and stationery that had been provided to girls. In some instances, boys stole learning materials from girls, or harassed them at school.

Save the Children (Ethiopia) midline evaluation report quotes a caregiver:

> *[…] it is only the girls who receive the school materials […] our boys feel excluded or not given the appropriate things they need to school. As a result our boys tend to take some school items from girls who receive school materials using force. This puts them to fight for the limited resource distribute to them and I believe this should be done for both kids.*
>
> (Caregiver in Ada'ar, Save the Children's midline evaluation report, p. 42)

ChildHope (Ethiopia) in their midline evaluation report explain that "the possible negative effect of the project on boys, and on gender relations, was a consistent theme" emerging from their qualitative research. Even though boys were being engaged through "good brother" clubs where they were taught about gender equality and the importance of supporting girls to get an education, there was evidence of boys resenting girls and feeling discriminated against by the GEC, of boys harassing girls and stealing from them, and even of boys dressing as girls to receive GEC education materials.

> *Boys in our class always get annoyed when we receive the educational materials. Sometimes they beat us and try to snatch our exercise books. Usually, they steal books from our bags. There is a boy in my class who always waits for us and asks us to give him our pens or exercise books otherwise he beats us.*

There was an older boy who has now dropped out and gone to Addis Ababa. He was insulting us. He said that we are better in class only because of the support of CHADET. He always was beating me out of class. I think he expected an exercise book from me. I know he took one exercise book from another girl.

Adults accompany their daughters to school when school materials are offered, to protect them from boys who threaten to rob them because they are disappointed they are not provided with them.

(Girl respondents, ChildHope Ethiopia's midline evaluation report, p. 43)

In CARE (Somalia) and Relief International (Somalia) project areas, community members stressed the importance of targeting boys as well as girls, while highlighting barriers to boys being in school and learning.

You cause discrimination. You prefer girls over boys, you improve girls only. Would you change the situation and help both boys and girls? Because boys need to be motivated.

(Father in South Central, RI midline evaluation report, p. 63f)

If you only train girls, you neglect boys. You say previously men discriminated against women. Just implement justice and help both girls and boys.

(Father from Puntland, RI midline evaluation report, p. 63f)

The CARE midline evaluation report states that the situation of boys is a "growing concern", as 8 per cent of boys dropped out during the current academic year due to poverty, lack of potential opportunities and employment in local communities, drug abuse, and overseas migration. The report concludes:

While gender disparities in favour of boys persist in enrolment, urgent action is required to prevent an inversion of the current situation, as already observed in other fragile environments where boys become engaged in drug abuse, migration and criminal activities. An active investment should be made in analysing the needs of boys to inform future activities (funded under SOMGEP or other programs) to support these as well as girls.

(CARE midline evaluation report, p.96)

WUSC (Kenya) have observed a "backlash" from communities, which feel that the GEC is only helping girls. Communities have intensified demands for greater equality and focus on boys. The project responded by producing informational and advocacy products that explain how the projects benefits both boys and girls and why it is worth educating girls (as well as boys); engaging men and boys and champions for girls' education and providing some learning materials to boys. Nevertheless, the WUSC midline evaluation report describes that perception persists in some communities, especially over the provision of solar lamps and remedial classes to girls but not boys. The project midline evaluation report concludes:

An important lesson is that boys can be as marginalised as girls in certain contexts and a sole focus on girls in a context of extreme poverty can result in this type of negative backlash. In the future, positioning the projects within communities where such extreme poverty, competition for scarce resources, and highly restricted life chances exist, will need to be carefully communicated. Interventions will need to target girls while not excluding boys.

(WUSC midline evaluation report, p. 68)

Save the Children (Mozambique) have also faced issues with backlash from the community and have in response changed their delivery approach. The project distributed 20 per cent of planned school kits to boys and integrated a small proportion of boys in girls' club activities. The project reports this to be an "effective strategy to avoid backlash and acknowledge the vulnerability of boys" (Save the Children midline evaluation report, p. 58).

In its evaluation report, Plan (Sierra-Leone) also refers to a recent study by Street Child which highlights the resentment from boys about the level of investment and encouragement that girls receive related to education. The report states that "going into the final year of activity the project will work where possible to target indirect beneficiaries to ensure there is no resentment towards GEC participants". However, the report does not clarify how this will be implemented.

Targeting and inequality

For some interventions, projects selectively targeted the most marginalised girls within their communities rather than providing universal support to all girls. This led to disappointment among some girls and families who did not receive any support and felt left out.

> *Relatives whose children were not chosen by "Vas-Y Fille" ask many questions like, "Why did the project not choose my daughter?*
>
> (School staff, Bandundu, IRC project area)
>
> *When Save the Children gave bags to half of the students there was unhappiness and complaints from parents whose children did not receive a bag. They are still hoping they would bring more. This kind of aid is not good for the psychology of the other students who did not receive it.*
>
> (Caregiver, Afar, STC project area)

Some respondents warned of the adverse incentives that the targeting of the most marginalised girls might create:

> *If they support only excluded students, then other students will also leave the school to become excluded from the school to receive the financial assistance.*
>
> (Community Leader, Nangarhar, BRAC project area)

Respondents also reported parents who had transferred their girls from one school to another, after learning that GEC activities were happening there. This might also explain some of the spill over effects observed in the SCW control areas.

Increased demand for education is not met by increased supply

There is some evidence that awareness raising activities and other support provided through SCW interventions (as well as by other non-GEC actors) have led to more girls enrolling in school or returning after a period of absence, for example in CARE and Relief International project areas in Somalia. Evidence from our quantitative analysis of barriers such as classroom overcrowding and IDIs shows that this increase in demand was not always met by an increase in the supply of education opportunities. Especially in Somali project areas where average enrolment increased by more than 10 percentage points in the treatment group, classrooms have become more crowded and teacher shortages have become even more acute.

> *On the sides of the classes, when the Relief International agency started to pay the fees of girls' school, it happened that more girls have been enrolled in the school and that the class became overcrowded, so we couldn't divide them into two classes due to that we couldn't cover the costs of two teachers.*
>
> (School staff, Puntland, CARE project area)
>
> *When [families] hear that schools are free, they fill it. This is […] children, who cannot get a proper education at home, and cannot afford expensive school, so he comes to this free school. And teachers cannot teach them well because of low pay. The students are now tremendously increasing.*
>
> (School staff, Somaliland, Relief International project area)

Promoting ambiguous messages about roles, norms and lifestyles

The EM's analysis of IDIs suggest that some of the messages conveyed by girl clubs or community mobilisation campaigns may be ambiguous and affect the girls' or communities' identity in ways that were not primarily intended. For example, some respondents valued girls' clubs not for giving girls a voice or making them more empowered, but for teaching them "respectable" behaviour in ways that may actually perpetuate gender stereotypes and the differential treatment of boys and girls. While this was only observed in some places, examples include:

> *They are taught how to behave.*

(School staff, Matabeleland South, World Vision project area)

They work on the hygiene of the class and its paint. They also aware themselves to wear decent clothes and not to wear makeup. And also to wear socks. They are doing my job and that is something amazing.

(School staff, Afar, Save the Children project area)

From IDIs, we find patterns of community members turning against their own livelihoods and cultural heritage as a result of their changing attitudes towards girls' education.

It changed their belief about education and they realized that being a pastoralist is not important and education can bring change.

(Caregiver, Afar, Save the Children project area)

We meet with them every 15 days or every month and we give them advice not to be too conservative and to allow their daughters to go to school, so that they can acquire a modern education which enables them to lead a modern life than to remain in this dry land rearing goats.

(School staff, Afar, Save the Children project area)

Now a life limited to following cattle steps is no longer wanted here. We saw that the animals are useless and whenever the rain stops to fall, they will die. Now we understood that education is better.

(Community leader, Afar, Save the Children project area)

In the thematic research, we found that in contexts of economic stress and climate change, a daughter's schooling can be an important element in her parents' long-term livelihood diversification strategy, because of the potential jobs she might be able to access with a school qualification. Diversifying income streams can enhance the parental household's resilience in the face of further shocks and stresses. This,is particularly the case for pastoralists whose traditional livelihood activities are directly threatened by climate variability and other factors, such as the Afar pastoralists in Ethiopia (Save the Children project area) and Samburu pastoralists in Kenya (CfBT project area). In Samburu East, we found from the thematic research that this view is strengthened by an official discourse which encourages pastoralists to move into more 'modern' livelihoods involving a more settled and sedentary way of life:

Parents have come to realise that educated kids can get jobs. If all they can do is herd, they are vulnerable to droughts; everything you have can be wiped out. So parents have seen the benefits of schooling.

(Girl respondent, Kenya, EM Thematic Research, Narrow Windows, Revolving Doors)

There appears to be a tension in some contexts between desires to diversify household livelihoods by means of educating girls on the one hand, and respecting and maintaining cultural integrity and tradition on the other.

Increased time poverty among beneficiary girls

In the EM's qualitative interviews many respondents suggest that awareness raising campaigns and the provision of incentives (e.g. solar lamps or bicycles) have changed attitudes, encouraged families to send girls to school, and raised their aspirations for girls' education. However, this also means that families expect more from girls than before, as they have to reconcile their school work with household chores and livelihood activities. One community leader emphasises that such duties remain important, given the prevailing high levels of poverty:

The time to weed, she knows […] is 08:00 am. By the time she leaves the field to reach home, [and] takes a bath, she will be late. Reaching school, it is whip. […] I think that […] IRC tries a little to organize training to tell parents to reduce work. Now, reducing work, how do you reduce it because it is in these fields that there is a life. There is no other work here in our home to say that I remove a little to pay for the school fees. It is only the field, if you do not cultivate there will be no studies.

(Community leader, IRC project area, DRC)

The pressure on girls to work is high in Somalia, where women and girls are often the families' main breadwinners.[63] Despite a change in awareness and attitudes towards girls' education, girls and women still have to carry out household chores and livelihood activities:

63 Judith Gardner and Judy El-Bushra (2016), The impact of war on Somali men and its effects on the family, women and children" Rift Valley Institute, Briefing paper.

The other issue that I would like to raise is that most of the family breadwinners are women, and most of the men are unemployed. [...] My wife and her daughters work for our family and they are the ones who make the living. I mostly do elderly tasks and mainly busy for social related issues […]. God blessed Somalia with their women's efficiency, as they were the ones who have been working for their families since we returned from the civil wars.

(Caregiver, CARE project area, Somalia)

Summary of key findings – Unintended consequences of SCW activities in target communities

The problem of boys resenting the prioritisation of girls emerged from many project midline evaluation reports, in a variety of countries and regions. To some extent, resentment can be expected as the privilege of one group is being reduced to the benefit of another. It may also reflect a masculine sense of entitlement. However, our analysis on enrolment and learning outcomes supports the notion that boys are not significantly less marginalised from education than girls in many SCW project areas – at least at primary level. As highlighted by WUSC (Kenya), prioritising one group over another in a context where resources are extremely scarce and where all children are struggling may create backlash and "interventions will need to target girls while not excluding boys".

Positive shift in attitudes towards educating girls is sometimes linked to a desire to diversify household livelihoods they may also convey the desire to overcome current livelihoods that are perceived as precarious. In certain contexts, such as in pastoralist communities, the association of girls' education with a modernisation narrative is related to tensions between cultural tradition and individual rights-based approaches to both education and development more widely. This is a tension that SCW projects should be critically aware of.

Key lessons learned

- Our findings about the resentment of the prioritisation of girls underscores the importance of selecting target areas carefully following **comprehensive gender analysis of local gender and power dynamics**, to ensure that girls are only prioritised where they are significantly disadvantaged with regards to boys. Interventions such as teacher training, the construction of schools, and the provision of learning materials may be more helpfully pitched as a universal intervention to improve education, compared to interventions that are specific to improving girls' education outcomes compared to boys.
- Gender analysis at the design stage should be **grounded in a thorough situation analysis of the cultural and social context in which each SCW project will be operating**. It is crucial to ensure that GEC interventions "do no harm" and that activities do not lead to perverse or adverse effects, as well as to allow projects to anticipate and mitigate such effects when they arise. Pre-baseline analysis should be followed by **close monitoring of the potential unintended impact of different activities on the life of girls and their communities** to ensure that the potential frictions between the messages and effects promoted by the GEC and the local lifestyles and livelihoods are properly managed and accounted for. Where tensions and misunderstandings are detected, projects should explore their causes (drawing on gender expertise where possible) and appraise possible mitigation strategies that can be put in place at any stage in the project's lifecycle.

http://www.politicalsettlements.org/files/2016/02/The-impact-of-war-on-Somali-men-and-its-effects-on-the-family-women-and-children-by-Judith-Gardner-and-Judy-El-Bushra-RVI-Rift-Valley-Forum-Briefing-Paper-2016-1.pdf

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

Sustainability was defined at the programme level of the GEC through two programme logframe outcome indicators:

- **Outcome 3: Match Funding**[64] – *Additional match funding secured alongside DFID GEC funds*.
- **Outcome 4: Sustainability – *Number of projects that have mechanisms in place to enable marginalised girls to complete a full cycle of education.***

All projects include outcome indicators in their logframes that are defined in more or less the same way as the programme level indicators.

This section relies exclusively on project midline evaluation reports and FM match funding data. In **project midline evaluation reports**, projects were required to respond to the following questions:

- What is the project's sustainability strategy?
- To what extent has the project identified the pre-conditions for scaling up and /or sustaining its activities and results?
- How has the project strategically engaged with other organisations to achieve complementary effects?
- To what extent has the project leveraged additional investment?
- What are your plans for delivering sustainable results?
- What are the lessons learned about the scalability and sustainability of the activities delivered?

Generally, projects followed the above reporting structure. Project midline evaluation reports provided evidence of the different ways in which projects have engaged different stakeholders, in particular government and communities, in the design and delivery of project activities. Overall though, there was little explanation of how and why different engagement strategies would lead to the sustainability of specific activities or the outcomes achieved by individual beneficiaries. It is possible that this reflects a reporting deficit i.e. external evaluators and projects have not been able to sufficiently report and explain project sustainability strategies and their progress to date.

Summary of baseline findings

At baseline, we concluded that there was very little if any baseline evidence relating to the sustainability of the projects' interventions with little or no discussion of the implications of the baseline findings for the projects' sustainability plans. The report recommended that projects should consider these implications and may need to conduct further research to revise and strengthen their sustainability plans.

3.5.1 To what extent has the SCW leveraged additional investment?

The FM in the latest GEC Annual Report (July 2016) reports that the SCW has exceeded its midline match funding target of £8.1m to achieve match funding of £12.5m[65], which represents a 154 per cent over-achievement. At the end of March 2016, £125.6m of funding had been disbursed to SCW projects supporting 449,300 girls with improved learning – for every **£1 of DFID expenditure, SCW projects have achieved 10 pence of match funding (a ratio of 1:0.1)**. By comparison:

- GEC IW projects (£28.4m disbursed /179,900 girls improved learning) achieved match funding of £4.8m (a ratio of 1:0.17); and
- GEC SPW projects (£20.6m disbursed /110,010 girls improved learning) achieved match funding of £25m (a ratio of 1:1.21).

While interesting, these are not like-for-like comparisons. IW projects are delivering smaller-scale projects with budgets up to £2m and the SPW projects are partnering with four private sector partners: Coca Cola; Avanti;

[64] The GEC Grant Handbook defines eligible match funding as funding that is: (1) *In place* – funds are held in an account in the Recipient organisation's name; or (2) *Committed* – funds are committed to the Recipient irrevocably for this project and are expected to be forthcoming for the subsequent years of the project; or (3) *Free of obligation* - funds without obligation to any third party.
[65] Reported to the Fund Manager through project Quarterly Expenditure Reports (QERs)

Discovery Channel; and Ericsson who are required to co-fund (through cash and in-kind contributions) their DFID funding of up to £15m.

The latest FM's programme management information about match funding[66] includes data for 11 out of 14 projects. FM match funding data was not available for BRAC (Afghanistan), Save the Children (Ethiopia) and RI (Somalia). GEC projects are set a match funding target that they are required to achieve by the end of the project. At midline, the IW achieved nearly £2m match funding in Year 1, £3.3m in Year 2 and £6.6m in Year 3 to March 2016.

On average, the 11 SCW projects reporting match funding have raised £1.07m per project and achieved 73 per cent of their targets. The total amounts of match funding raised ranges from £19,473 (5% of lifetime target) by ACTED (Afghanistan) to £4m (81% of lifetime target) raised by Camfed (Tanzania and Zimbabwe).

Many projects did not refer to their reported match funding in their project midline evaluation reports. There are also several discrepancies between data provided in the reports and the FM data. This is most likely caused by a time lag between the two reporting processes. The following sections are based on a review of the information provided in the project midline evaluation reports.

At the programme level, it is not clear how match funding contributes to or relates to the sustainability of project activities. Match funding provides an indication that stakeholders have bought into the concept, aims and objectives of projects. But this does not necessarily mean that match funding is available to sustain project activities beyond the life of the project. This could have been explained and evidenced as an integral part of the sustainability plan or strategy. Generally projects did not provide much explanation or evidence of how match funding (cash or in-kind) contributed to the sustainability of the project. Notable exceptions were:

- CfBT (Kenya) chose to partner with the Teacher Service Commission (TSC) in Kenya to support its teacher training and teacher coaching initiative. This partnership resulted in the TSC providing 70 per cent of the teacher coaches working on the project and paying their full salaries. The project provides logistical support to these teachers. More importantly from a sustainability perspective, it has enabled the project's teacher coaching to be built into an existing government structure. However, the sustainability of this part of the project depends on whether the TSC will continue to allocate resources supporting the teacher coaching. CfBT's midline evaluation report comments that the project would have benefited from cooperating earlier and more intensively with the TSC rather than the Ministry of Education. CfBT is one of the few projects to pay considerable attention to its sustainability plan – in Annex 8 of its midline evaluation report, it includes a detailed breakdown of the resources it has levered into the project.
- Aga Khan Foundation (Afghanistan) reports a contribution of £736,787 by communities providing classroom space and £192,932 raised through volunteer time invested by those people serving on school management councils and those supporting AKF's education initiatives. AKF's midline evaluation report also includes detailed unit cost analysis of the provision of Community-Based Education (CBE) classes compared with the provision of education through government schools. The report comments that the greater cost per child of AKF's CBE classes is offset by the benefits realised through improvements in gender equity and retention rates in CBE classes. However, the report comments that the Ministry of Education perceives CBE as a short-term stop gap solution provided by NGOs rather than a long-term solution that the government should adopt. This is a challenge that BRAC in Afghanistan also faces in delivering its CBE.

3.5.2 To what extent has the SCW influenced changes in the behaviour and practice of others?

All projects acknowledge in their midline evaluation reports the importance of **engaging with government** to stimulate and facilitate buy-in at national, provincial, district and local level as a pre-condition to sustaining key project activities. Several projects are able to demonstrate that they have engaged with government in a variety of ways, particularly by:

[66] FM GEC Project Match Funding programme management data 22 November 2016.

- aligning their project designs and interventions with national government policies, education sector plans – for example, BRAC (Afghanistan) has consistently followed Community-Based Education (CBE) government policy guidelines in the design and delivery of its community-based girls' schools;
- actively participating in Ministry of Education (MoE) working groups, steering committees and policy fora at different institutional levels to raise awareness of what projects are doing and influence policy-making;
- involving government in the design and delivery of the project through regular collaboration with staff district and local government education offices;
- developing the capacity of local government staff to adopt different types of approaches used by projects – for example, RI (Somalia) is building the capacity of the MoE by providing training in leadership and gender sensitive education, as well as monitoring against quality assurance mechanisms; and
- using information and evidence from evaluation, research and progress reports to shape and influence key policy decisions – for example, AKF (Afghanistan) has inputted into the National Education Strategy Plan, and CBE Policy, and Camfed (Tanzania) is supporting the development of MoEST guidelines for new government education policy to abolish school fees at junior secondary school level.

Several projects have reported the difficulties that they have experienced in collaborating with government and trying to get enough buy-in and ownership to sustain key activities beyond the life of the project – for example:

- Care (Somalia) stresses that government capacity is too low to expect it to sustain the project's activities – this problem is further complicated by the need to work with three Ministries of Education across Somalia. As a result, the Care has invested in developing 'grassroots capacity' by training Community Education Committees in education management;
- WUSC (Kenya) highlights the difficulties in working closely with the Teacher Service Commission (TSC) at a local level, while key decisions about teacher recruitment are made centrally in their office in Nairobi who WUSC engages with less frequently;
- RI (Somalia) flags high staff turnover in the MoE as a key factor limiting the sustainability of the project's impact and getting the buy-in of key staff; and
- CfBT (Kenya), WUSC (Kenya), ACTED (Afghanistan), AKF (Afghanistan), BRAC (Afghanistan), Care (Somalia), RI (Somalia) all flag a lack of government resources and budget as a significant constraint to the sustainability of key project activities. For example, AKF's (Afghanistan) midline evaluation report highlights 'a persistent lack of funding and a general lack of interest in sustaining CBE classes on the part of the MoE'.

A lack of government funding and capacity more generally is a fundamentally important constraint to any improvements in government education systems. It is particularly important if sustainability strategies are highly dependent on further MoE funding or MoE taking over project activities, schools, bursaries or teachers at the end of the project. Projects are very unlikely to be able to directly address the problem of a severe lack of government resources. In certain situations this may mean that the only realistic sustainability strategy for critical investments like schools, classrooms, teacher salaries etc. is the need for the project to find further external funding. However, only a few projects explicitly recognised this in their reports.

While projects demonstrated in their reports that they have actively engaged government in different ways, it was generally not clear what projects' strategies were for influencing a change in government behaviour and with what specific effects on the sustainability of different parts of the project. An exception to this, was CfBT (Kenya) which was able to describe in detail in their midline evaluation report the different ways in which their engagement with government education bodies was built into the design and delivery of their interventions.

Several project midline evaluation reports suggest that they are at a relatively early stage in their engagement with government as part of their sustainability strategies. For example, ChildHope (Ethiopia) sustainability plan started in January 2016, with several critical activities starting in the final year of the project. Many projects describe a focus or refocus in the final year of implementation on advocating to government to take over various project activities. Many projects seem to be focusing on developing and delivering their sustainability strategies too late in the delivery process.

In their midline evaluation reports, projects predominantly discuss the different ways in which they have **engaged with communities** and the contributions local people and community groups have made in driving and delivering project activities to date. This level of community engagement across the SCW provides a potential platform for sustaining project activities. SCW projects though were not able to explain how and why communities will continue to sustain particular activities once the project ends.

Generally, midline evaluation reports include ambiguous assertions that community buy-in, engagement and involvement to date will translate into activities being adopted, driven and sustained by communities themselves after projects have finished. This is unlikely to be an effective sustainability strategy for activities that need to continue to be delivered, especially for voluntary groups that are still relatively new or new activities that have not yet established a robust structure to support them. Reports did not describe or explain what formal or informal infrastructure was available to continue supporting voluntary groups to sustain project activities. Critical mass and attitudinal change can certainly lead to the behaviour change required to sustain new types of educational support activities. But projects' midline evaluation reports contained little detail or explanation about how the conditions required for sustainability would be established by the end of the SCW in July 2017.

In the absence of government capacity and funding, many of the SCW projects' sustainability strategies appear dependent on continued community support. However, it is unclear how and why community-based support will succeed in sustaining the level of involvement and contributions needed to continue to deliver the same results.

3.5.3 Have projects strategically engaged with other organisations to achieve complementary effects?

A key finding from Section 3.1 is that the environments that SCW are operating in are very crowded spaces. We found extensive evidence of non-GEC education activities happening across both treatment and control areas. The project midline evaluation reports provide little evidence of joining up with other non-GEC activities to achieve potential synergies and to avoid duplication or at worst conflicting and counter-productive activities (see the example of CfBT in Box 4, Section 3.1).

3.5.4 To what extent have SCW projects identified the pre-conditions for scaling up or sustaining their activities?

The FM in its GEC Annual Report (July 2016) reports that all 14 of the SCW projects have mechanisms in place to enable marginalised girls to complete a full cycle of education. According to this assessment the SCW has achieved and exceeded the sustainability targets set out in projects' logframes. However, in the Annual Report, the FM recommends that a better indicator is required for Outcome 4 to better capture the extent to which the projects in the GEC are sustainable at different stages in their lifecycle.

At midline, across the SCW there is little reported evidence that projects have sufficiently considered from the start what conditions they need to develop to sustain the benefits for marginalised girls. There is a distinct lack of evidence of analysis of the strategic landscape with regards to sustainability. For several projects, strategic engagement for the purpose of sustaining project activities appears to have happened relatively late in the implementation period. CfBT (Kenya) is a clear exception as the project purposefully set out to align with and complement its strategic landscape as an integral part of its project design. BRAC (Afghanistan) and AKF (Afghanistan) from the start based their sustainability strategies on transitioning their activities over to government. However, both BRAC and AKF report that transition to government remains uncertain and a significant challenge. For example, AKF reports[67]:

> *A persistent lack of funding and a general lack of interest in sustaining CBE classes on the part of the Ministry of Education constitute a significant risk to the assumption that the Ministry of Education will either ensure that there are sufficient resources for students to transition to government hub school or for CBE classes to be taken over by the government. There has been recent progress in discussions on the handover of students in the CBE classes funded by the Girls' Education Challenge in Afghanistan. NGO partners, donors and MoE officials participated in a number of meetings in winter 2015 to discuss the issue of handover and the need to jointly plan to ensure a smooth transition of project. This will lead to a common effort to revise National Education Strategic Plan (NESP III) by the end of 2016. However, barring provision of funding specifically earmarked for sustaining CBE classes and a firm commitment by the Ministry of Education to support classes established with off budget funding from donors, it seems unlikely that significant numbers of students will be able to continue*

[67] Corboz, J. Cottereau, S. (2016) Midline GEC Report – Steps Towards Afghan Girls' Educational Success: AKF UK 5147 (pp.78)

their education and it seems particularly likely that girls will be especially affected by the government's inability to support classes and students.

A key pre-condition to sustainability is knowing which activities or interventions are the most effective (and cost-effective) in delivering an impact for marginalised girls and as a result are critical to sustaining the delivery of long-term benefits. As discussed in previous sections, at the moment, it is difficult to evaluate and identify which interventions are proving most or least effective and it is also unclear in projects' midline evaluation reports. Until projects are able to identify what should and should not be sustained it will be very difficult for them to develop effective sustainability strategies.

Summary of key findings – Sustainability

The SCW exceeded its midline match funding target of £8.1m to achieve match funding of £12.5m, which represents a 154 percent over-achievement. It is unclear how much of this match funding indicates initial stakeholder buy-in to project approaches at the start and /or a commitment by stakeholders to sustain project activities.

Several projects demonstrated that they have engaged with government in a variety of ways. However, many projects reported the difficulties that they experienced in collaborating with government and trying to get enough buy-in and ownership to sustain key activities beyond the life of the project. In the absence of government capacity and funding, many of the SCW projects' sustainability strategies appear dependent on continued community support. It is generally unclear how and why community-based support will succeed in sustaining the level of involvement and contributions needed to continue to deliver activities capable of achieving similar or better results.

Overall, there is little evidence that the SCW will deliver sustainable results by the end of the programme in July 2017. For many projects it also appears that they have left it too late to successfully deliver their sustainability strategies.

3.6 To what extent does the SCW represent good value for money?

3.6.1 Approach and project response

Our approach to assessing the value for money of the SCW at the midline stage, two years into the three year project implementation period, is to:

1. **Use our midline findings from this window-level impact evaluation** to determine the value generated to date by SCW projects. These findings were informed by the **Fund Manager's Metrics Report**[68] for the GEC produced in November 2015. This report uses VfM metric tables created in 2013 as a tool for calculating economy VfM indicators for Years 1 and 2 of the GEC. We have used and referred to these previously in this report to estimate the proportion of budgets spent on different types of interventions.
2. Use the evidence of value for money presented in the **SCW project midline evaluation reports**. Projects were provided with the following structure and guidance[69] about how to evaluate and report their value for money at midline:
    - *Describe and evidence the extent to which key decisions and options that were considered about the design and delivery of the project have explicitly driven decisions about the scale and types of costs that have been incurred to date to ensure that best value for money is achieved [cost drivers].*
    - *To what extent has the project been efficient in delivering its outputs?*
    - *What types of interventions are proving most cost-effective in terms of their effects on the project's outcomes, for whom and under what types of conditions?*

All SCW projects included a section on value for money in their project midline evaluation reports except ACTED (Afghanistan). ACTED's project has now come to an end. So, their report was effectively an Endline Evaluation Report. The following sections draw heavily on the value for money sections of the projects' midline evaluation

[68] PwC (2015), 'Value for money (VfM) metrics on the Girls' Education Challenge – Report on VfM expenditure metrics for the GEC Year 1 and 2: January 2013 – March 2015'.
[69] Fund Manager (July 2015), 'Project Midline Evaluation Report Template' p.11.

reports. On the whole these sections were generally poor. They provided little information and evidence that could be used to draw meaningful or conclusive findings at the window level about cost drivers, efficiency and cost-effectiveness.

3.6.2 What design and operational factors have driven decisions about the costs that projects have incurred?

ChildHope (Ethiopia) was one of the few projects to make a link between evidence-based decisions about their project design and the ways in which they allocated their budget. Very few projects were able to adequately report how and why decisions about their project design drove their cost structure, for example by outlining what type of interventions represented the biggest investment and why this was justified. Care (Somalia) highlighted the higher costs involved in targeting marginalised girls in remote, contested and hard-to-reach areas and that school construction represented their biggest cost driver, but there was little more than this across the SCW project midline evaluation reports.

Several SCW projects were able to expand on decisions that they made about how to deliver particular interventions, and how and why cost played a role in these decisions. Care (Somalia) reports that it decided to use a clustering method to reduce the logistical costs of delivering its activities to communities and schools in the same geographic proximity. BRAC (Afghanistan) reported using local volunteers 'at no cost' to the project to deliver its monthly mothers' forum. However, this assessment did not to acknowledge that voluntary support like this comes at a considerable cost to local people in terms of their time and effort, which needs to be factored into any sustainability strategy and value for money assessment.

Box 14: Camfed's cost-effective approach to the design of the project's Step Up Fund

Camfed's Step Up Fund (SUF), currently being implemented in Tanzania and Zimbabwe is a good example of how cost has been considered in the way that it has been design and delivered. The fund acts as a low-cost alternative to a more comprehensive bursary package. It is a targeted response mechanism that has been designed to meet the school-going needs of marginalised girls. The SUF is decentralised and is being implemented and managed through existing local structures to minimise the 'leakage' of resources due to administration costs. The report states that the SUF is proving effective and that Camfed is undertaking additional research in the final year of the project to 'assess the extent of the community buy-in and engagement in the delivery of the SUF and the types of conditions under which this is incentivised. This will have important implications for the scalability and replicability of the SUF approach, particularly with a view to possible future take-up and replication by Ministries'.

This has been highlighted as an example of good practice, because it demonstrates consideration of cost-effectiveness at the design stage that is being evaluated as a stand-alone intervention to assess its impact, effectiveness and cost-effectiveness.

The above examples demonstrate some awareness of how costs influence and inform project design and delivery decisions. There are however only a few examples like this across the SCW. Furthermore, even these examples do not progress to the next step to explain and evidence the extent to which these key decisions enable the project to deliver effective and impactful interventions.

3.6.3 What types of interventions are proving most cost-effective, for whom and under what types of conditions?

Apart from the Camfed example provided above, there is very little further evidence of cost-effectiveness. A key constraint that projects face, which was highlighted in our analysis of the effectiveness of the SCW (Section 3.4) is not being able to assess which interventions were the most effective and impactful. If projects are unable to evaluate, even qualitatively, which interventions are having the greatest effect, on whom, why and under what types of conditions then it becomes very difficult to assess which types of activities have proved to be more cost effective than others.

Box 15: Key lesson learned about measuring cost-effectiveness of the SCW

As part of the GEC application process a key criterion for assessing project applications was the cost per girl benefiting from the project i.e. an outcome unit cost for each marginalised girl with improved learning. The SCW supports marginalised girls across very diverse contexts. The extent to which SCW girls are marginalised compared to other girls and boys in their communities varies from one project context to another, and from one community to

the next. The barriers that marginalise girls from education are very context-specific. That means that any given type of input cost that is needed to address a particular barriers to girls' educational marginalisation could vary significantly from one context to another.

The cost of intervening in any given context will be driven by factors including: the local prices of inputs; logistical difficulties in reaching girls and their communities; and the complexity and severity of the barriers that need to be overcome for a marginalised girl to achieve a measurable improvement in their literacy and numeracy. For these reasons, any cost-effectiveness comparisons of the unit costs for each girl with improved learning between projects would be undermined these contextual differences. Similarly, an aggregate outcome unit cost for the cost per girl with improved learning would also be of little use because a sufficiently similar external benchmark could not be found due to the variety of contexts and cost structures that SCW projects operate within.

Cost-effectiveness analysis could have been useful by comparing the outcome unit costs at midline with those at baseline for each individual project. In effect this means **each project using its own baseline outcome unit cost as a benchmark for measuring changes in its own cost-effectiveness at midline and endline.** Tracking changes in outcome unit costs for a project and identifying the reasons for these changes could help explain how and why a project make more or less cost-effective over time as an integral part of the evaluation process. Project M&E Frameworks though have not been set up to track outcomes and costs in this way, which is a valuable lesson learned.

3.6.4 To what extent have SCW projects been efficient and economical in delivering their outputs?

Almost all midline evaluation reports mentioned the procurement and management practices and operational strategies that they had in place to achieve the best price for inputs of to reduce costs. These included:

- bulk buying /combined purchasing to achieve economies of scale;
- competitive procurement processes to buy capital equipment and consultancy services;
- benchmarking fees and wages;
- working through a consortium – including: using centralised management to reduce management and administration /overhead costs; benefiting from a larger geographic footprint; exchanging ideas and information;
- cost sharing by using existing offices and staff;
- using local staff; and
- using the people from the local community /PTAs as community mobilisers /delivery agents /local labour to construct classrooms – also as a strategy for obtaining community buy-in and to enhance potential sustainability.

All of the above practices and strategies represent good practice to achieve the most economical inputs possible and to keep the cost of activities down. Very few projects though were able to provide any cost data or evidence to show what effect these practices had on their costs in absolute or relative terms e.g. cost savings as a percentage of the total project cost or of the cost of each type of intervention.

There was very little evidence reported across the SCW of the efficiency of delivering outputs. AKF (Afghanistan) did conduct a comprehensive unit cost analysis of providing education to each student attending its Community-Based Education (CBE) classes. The AKF midline evaluation reports stated that: 'While government schools are spending less per student per year as compared to CBE classes, it is spending significantly more to keep a child in school until grade four, because so many students, especially girls, are not retained'. This analysis by AKF is a credible start to considering the lifetime benefits of this type of education provision, which include greater student retention compared to the cost of higher drop-out rates in government schools. To strengthen this analysis, analysis of the full lifetime benefits and costs needs to also take into account the challenge that all the GEC projects in Afghanistan face in convincing the Ministry of Education to take over or adopt the management and administration of their schools and the costs associated with this transition.

3.6.5 Summary of key findings from the impact evaluation of the SCW

The value for money of the SCW as a whole at midline should be informed by:

- education impacts to date, which is the ultimate objective of SCW projects and the value that projects add;
- projects' effectiveness; and
- the potential sustainability in terms of the activities they have delivered and the persistence of the benefits realised by the girls that the projects have targeted and reached.

Box 14 below provides a short summary of the key findings presented in previous sections. The purpose of this summary is to collate in one place the most relevant headline findings that help inform an interim judgement about the overall VfM of the SCW at the midline stage.

Box 16: Summary of key findings relevant to the midline VfM assessment of SCW

Reach and equity:

- We cannot systematically determine from our EM sample whether girls targeted by SCW projects are the most marginalised girls in their communities or regions. Girls in target communities are disadvantaged to varying degrees and across different dimensions. Social and economic markers of marginalisation differ substantially from one project's context to another and show the high level of heterogeneity among SCW target populations. No common definition of marginalisation was prescribed at the programme level and marginalisation has been defined differently by different projects. Most projects have used geographic criteria, while others used ethnographic targeting criteria. This makes it difficult to assess the extent to which projects effectively impacted on the barriers that marginalised their target girls compared to other girls and boys in the community.

Impact on literacy:

- EM data shows no evidence at the window level of a positive effect of SCW projects on the average reading fluency of girls in treatment communities compared to control communities. There is some variation at the project level with two projects showing significant improvements in literacy.
- Six out of fourteen projects met or exceeded their literacy outcome targets at midline according to their own impact data that was reanalysed by the FM.
- The SCW has had a positive statistically significant effect of four words per minute on the *median* reading fluency score. At baseline, across the EM sample more than half the children who were tested could not read a single word. Now, at midline, half the children can read at least 27 words per minute.
- Out-of-school girls are the only identified subgroup of marginalised girls that the SCW had a net positive effect on in terms of learning. Reading fluency improved significantly by 12 words per minute among out-of-school girls since baseline compared to the control group.
- Gender analysis from the DRC, Ethiopia and Kenya project areas confirms our baseline findings that in-school girls are not systematically disadvantaged compared with in-school boys of the same grades (P4 and P6) in these areas. The gender gaps related to being-in-school and learning were relatively small at baseline, at primary age. Girls in treatment and control areas showed greater improvements in literacy and numeracy scores than boys of the same age and school year. In other words, girls in this sample of primary school students are now overtaking boys across the SCW.

Impact on numeracy:

- EM data shows no evidence of a window-wide effect on the numeracy of girls in treatment communities compared to control communities. The EM data suggests that numeracy improved significantly in one project area.
- Similar to literacy, numeracy improved significantly for out-of-school girls by 13 percentage points since baseline compared to the control group according to the EM data.
- Three out of fourteen projects met or exceeded their numeracy outcome targets at midline according to their own impact data that was reanalysed by the FM. Three other projects partially achieved (i.e. greater than 60%) their numeracy targets.

Impact on attendance:

- EM data shows that attendance (and enrolment) rates have not significantly improved across SCW target communities compared to control areas. However, the difficulty of capturing small changes through household-based measures prevented us to produce conclusive findings at midline about improvements in girls' attendance at school.

Effectiveness of SCW projects' activities:

- Across the SCW the majority of the most important barriers to girls' education identified at baseline in target communities has not changed significantly as a result of SCW activities. Only a few barriers, mostly related to community attitudes and caregivers' satisfaction with school facilities, teaching and school affordability seem to have changed significantly across the 14 projects.
- The activities which seem to have had the biggest impact on girls' learning are special tutoring, help with school work and teacher training. Projects who invested the most in economic interventions are the same that managed to improve school affordability substantially. The provision of bursaries and in-kind support, as well as loans seems to have been an effective way of alleviating the burden of school-related expenses in some communities, but this does not seem to have translated into better learning outcomes yet.
- It is very difficult to assess the quality of teaching and quality of school infrastructure through quantitative surveys. However, what appears from our analysis is that the most successful projects are those who have impacted on a range of different barriers to education that meet the right needs of their target population in an effective way. This does not mean that those projects equally invested in a range of different activities, but rather that they were aware of the precise and specific needs of their target populations and invested strategically to address these needs.
- Several projects experienced and reported a backlash and resentment from boys that may partly be explained by boys being unwilling to give up gendered privileges, but it might also reflect that boys have many similar needs as girls.

Sustainability:

- The SCW exceeded its midline match funding target of £8.1m to achieve match funding of £12.5m, which represents a 154 per cent over-achievement. It is unclear how much of this match funding indicates initial stakeholder buy-in to project approaches at the start and /or a commitment by stakeholders to sustain project activities.
- Several projects demonstrated that they have engaged with government in a variety of ways. However, many projects reported the difficulties that they experienced in collaborating with government and trying to get enough buy-in and ownership to sustain key activities beyond the life of the project.
- In the absence of government capacity and funding, many of the SCW projects' sustainability strategies appear dependent on continued community support. Although it is generally unclear how and why community-based support will succeed in sustaining the level of involvement and contributions needed to continue to deliver the same results.
- Generally, there is little evidence at the midline stage to suggest that the SCW projects will be able to deliver sustainable results by the end of the programme in July 2017. This situation is likely to change. However, for many projects it also appears that they have left it too late to successfully deliver their sustainability strategies.

The above headline findings summarise what we generally found from the EM's primary research and the evidence reported by projects in their midline evaluation reports. Generalised findings like this mask variations and exceptions at the project level, with some projects achieving more than others in terms of their impact and the effectiveness of their interventions.

Summary of key findings – Value for money

Generally, SCW projects have targeted and reached girls who are marginalised from education in very different ways across diverse contexts and to varying degrees. At midline, the SCW has had a significant impact on out of school girls and those at the bottom end of the learning spectrum. However, overall, projects have had a relatively

limited impact on literacy and numeracy compared to the midline targets set for them at the outset. Their impact on attendance also looks low and both the project and EM data is inconclusive in this respect. Many barriers that were found to be important to girls' learning have not changed significantly at midline – the biggest impacts on girls' learning resulted from interventions focused on special tutoring, help with school work and teacher training.

Projects were generally unable to report on the economy, efficiency or cost-effectiveness of their interventions in their midline evaluation reports. SCW projects have levered in a significant amount of match funding, which seems to be primarily focused on helping projects deliver their current activities. At this stage, there is little available evidence to suggest than many project activities will be scaled up or sustained by the end of the programme in July 2017.

Projects submitted their midline evaluation reports at the end of the second year of implementation. With one year remaining, the evidence available at the midline stage of the SCW suggests that a majority of projects will not achieve their learning outcome targets by the end of the programme. To achieve these targets projects will need to demonstrate a significant improvement in their performance and impact with the remaining resources for the last 12 months of the programme.

4 Conclusions

The SCW did not have a significant impact on GEC outcomes at the window level at midline. We acknowledge that this masks over-performance and under-performance at the project level. However, a majority of projects did not achieve their midline literacy and numeracy targets.

On average across the SCW the EM midline evaluation finds no significant effects on learning, enrolment and attendance for girls in target treatment communities. While literacy, numeracy and enrolment have improved since baseline there is no significant difference between the changes measured in the treatment and control groups. At the project level, however, six projects exceeded their midline literacy target according to the projects' own data. Three out of fourteen projects met or exceeded their numeracy outcome targets at midline according to their own data. Three other projects partially achieved (i.e. greater than 60%) their numeracy targets. Enrolment and attendance rates have not improved significantly across the SCW. For the latter, this is inconclusive because of the difficulties experienced by projects and the EM in collecting reliable attendance data.

At the window level, the EM data show out-of-school girls are the only subgroup of marginalised girls that the SCW had a net positive effect on. Their literacy and numeracy scores increased significantly, over and above the change measured in the control group. This suggests that projects have performed better at improving the learning outcomes for girls who start from a very low base. At the same time, a project like Save the Children (Ethiopia) has struggled to achieve any measurable improvement in learning outcomes even though its target populations started from a very low base at baseline. The EM data suggest that basic conditions for education and learning need to be in place to enable projects to deliver improvements at a significant scale. For example, we found that 99 per cent of schools in Ethiopia's pastoralist Afar region did not have access to safe water and 97 per cent of schools did not have access to electricity all day.

Many projects established girls' clubs, tutoring, or alternative education facilities for out-of-school girls to help them prepare for re-integration into mainstream schools or to teach them basic numeracy, literacy, life and vocational skills. The positive results observed for out-of-school girls suggests that creating additional or alternative learning environments have been effective in the short-term. Pre-existing school and classroom environments facing engrained problems associated with poor teaching, teacher absenteeism, and poor school governance may be more difficult to change within the limited implementation time available. However, it may prove difficult for projects to sustain solutions that are developed outside established school systems, and to scale up these types of approaches with schools and communities who lack funding and resources.

Overall, the SCW did not have enough of an effect at midline on the barriers that had the greatest impact on girls' learning. Projects have adopted holistic approach involving multiple types of interventions tackling the barriers to education that marginalised girls face.

Most of the important barriers to girls' education have not changed significantly at the window level as a result of SCW activities. This is in line with our findings relating to the lack of impact on learning, enrolment and attendance. Most projects though had an effect on a few barriers, but they were either not the most important constraints to learning, or the size and scale of the change was not sufficiently large to make a significant difference overall.

It seems that teacher training has generally been effective, even in established schools. CfBT (Kenya) and the Aga Khan Foundation (Afghanistan) invested 41 per cent and 24 per cent of their budgets respectively on teacher training and recruitment. They are among the only projects that show a significant effect on learning according to the EM data. In addition, girls who received special tutoring and girls attending schools where teacher training has taken place showed significantly greater improvements in learning than girls who had not participated in these types of activities (both in treatment and control groups). The EM data did not show a similar effect for the provision of bursaries and school materials, or the supply of funding to schools. Economic interventions seem to have been effective in making school more affordable, but this did not necessarily translate into higher enrolment or learning.

All SCW projects used a "holistic approach" to address a wide range of personal, family, community, school and governance factors. Holistic approaches are intended to tackle multiple barriers affecting girls' education. However, it is not clear from the evidence how or why projects prioritised the types of barriers they focused on and invested varying amounts of their budgets to address. An evidence-based understanding of which barriers affect girls' education outcomes the most is not evident in many projects' midline evaluation reports. We recognise that across

such a diverse portfolio of projects there are many factors (related to both project design and implementation factors) that could potentially explain why projects have been effective or not. However, projects addressing many types of barriers that have not had an impact at midline and are not sufficiently certain of whether they are investing enough budget in the right barriers may be lacking the focus needed to impact on the biggest education constraints that their target groups face. It is difficult to conclude the extent to which this has influenced the effectiveness of projects without more evidence related to their implementation, but it is a possible contributing factor.

Furthermore, holistic approaches may be more effective when delivered through a broad alliance of local government and non-government stakeholders who divide efforts rather than duplicate them. It seems that Aga Khan Foundation (Afghanistan) for example, was successful in working with communities to establish community-based and alternative education facilitates and hire teachers for these facilities – this seems to be based on good knowledge of the local context and the most pressing education problems that their target girls and communities face.

At the start of the GEC, SCW projects targeted a variety of different populations with different needs and starting points. Projects made trade-offs between the number of girls they reach, the degree to which girls are marginalised and the extent to which they are able to deliver sustainable outcomes.

SCW projects are targeting very heterogeneous girls and communities. Projects interpreted the challenge of the universal education outcome targets they were set at the outset very differently, which has led to projects making trade-offs between the number of girls they reach, the nature of girls' marginalisation and their entry points for intervention.

Midline analysis of marginalisation shows that SCW target girls are disadvantaged across a variety of dimensions that differ significantly between project areas. There is great diversity in environments, populations and challenges. Projects were not explicitly required to systematically analyse and articulate to what extent and how marginalised girls differed from other girls in their target communities, what their specific education needs are, and what distance they must travel to improve their learning outcomes significantly. Additionally, most project midline reports did not establish clearly who received what intervention, for what specific purpose, and with what effect given the girls' initial baseline needs. This makes it difficult to compare the performance of projects across the window.

For those projects working with particularly marginalised groups such as disabled girls, young mothers, or pastoralist communities, achieving their learning targets may be more difficult than for those working with a population of in-school girls who are not more marginalised than the rest of their communities. It is clear from the EM data that projects' decision about who they target has involved making trade-offs between: the number of girls they need to reach and benefit; the levels of severity and complexity of problems that marginalise girls from education; and the extent to which underlying causes have been addressed to deliver sustainable outcomes in the longer-term.

In order to be effective, projects had to fully adapt their intervention designs and delivery strategies to the local context, which requires building contextual analysis firmly into the design and M&E process. This is crucial to understanding how different factors of marginalisation interact in a given context to marginalise girls from education. Yet this process of adapting to the local context and the effects of contextual factors on projects' success is not evident in project midline evaluation reports.

SCW project areas are a crowded space. There is a lack of evidence that projects have mapped their strategic landscapes to identify potential partners and the best entry points. As a result, there appears to be little coordination or complementarity between GEC projects and non-GEC actors.

SCW communities are a crowded space for education programming and our evidence suggests that non-GEC activities are taking place in most treatment and control areas that address similar educational barriers. This may have led to contamination of projects' control groups and as a result their impact evaluation findings. However, regardless, there is little evidence that projects have coordinated with other GEC projects and non-GEC actors for the purpose of joint programme planning, design or implementation. Only a few project midline evaluation reports explicitly mentioned the activities of other actors in GEC treatment areas even though the EM's qualitative data suggests that these were prevalent in most project areas.

A key learning point from midline is that projects would have benefitted from conducting more comprehensive strategic landscaping before the start of implementation. The SCW had no formal Inception Phase and there was neither much time nor an explicit requirement for projects to map the strategic landscape, identify entry points from a political economy perspective, and scope strategic partnerships with government and other actors at the pre-baseline stage. While most projects did liaise with government to strengthen their implementation and sustainability, project evaluation reports typically showed little evidence of systematic political economy and contextual analysis. As a result, many project designs tackle a multitude of problems without demonstrating or explaining how project activities fit with and benefit from what others are doing in the same space.

SCW projects provided little analysis of how different types of interventions have addressed barriers that specifically impact on girls' education more than boys. As a result, the extent to which projects have improved gender disparity in learning and attendance outcomes at midline is not clear.

Our gender analysis of school-based data collected in project areas in DRC, Ethiopia and Kenya confirms our baseline findings that in-school girls of primary age (specifically P4 and P6) are not systematically disadvantaged compared with in-school boys of the same grades. Nevertheless, girls face particular gendered barriers, especially as they grow older and progress to secondary school.

The EM midline data suggests that girls in primary grades 4 and 6 do not have significantly lower learning outcomes than boys (see Section 3.3). It seems that key barriers such as poverty, poor teaching, inadequate school facilities and poor school governance universally affect learning conditions for boys and girls in primary school. Some of these factors, however, have a latent effect on girls more than boys that negatively impacts on their education as they get older and progress to secondary school[70].

Several projects experienced and reported a backlash and resentment from boys that may partly be explained by boys being unwilling to give up gendered privileges, but it might also reflect that boys have many similar needs as girls. These findings point to the importance of exploring gender gaps in context for the purpose of distinguishing between universal barriers to learning from gendered problems that affect girls more than boys. It is also critical to understand what effects gender dynamics have on girls' education immediately and how these evolve and take effect as girls grow older. This type of gender analysis was lacking in projects' midline evaluation reports. Testing boys in addition to girls was not a requirement for projects, who as a result, are not in a position to analyse where gender gaps exist and how they might have changed as a result of the SCW. Likewise, monitoring and evaluating changes in gender relations was not a standard component in projects' M&E frameworks.

A key learning point is that girls' education programmes like the GEC need to place gender analysis at the heart of the design process and M&E, in order to fully understand gender relations and intervene effectively without alienating men and boys in the process. Projects need to be able to appraise the risks of targeting girls over boys and develop relevant mitigation strategies from the start. They also need to consider whether there might be gendered, unintended consequences for girls such as having to juggle even more responsibilities and needing to meet even higher family expectations.

Sustainability needs to be locked into projects from the start and become an integral part of the problem diagnosis, project design, monitoring and evaluation processes to deliver best value for money.

The SCW exceeded its midline match funding target of £8.1m achieving match funding of £12.5m, but it is not clear how much of this match funding supports the sustainability of project activities. Several projects engaged with government in a variety of ways, but many struggled to generate enough buy-in and ownership to sustain key activities beyond the life of the project. In the absence of government capacity and funding, many of the SCW projects' sustainability strategies appear dependent on continued community support. It is generally unclear though how and why community-based support will succeed in sustaining the level of involvement and contributions needed to continue delivering the same results beyond the end of the programme in July 2017.

[70] The EM undertook thematic research[70] to explore why some girls aged 12-15 stay in school, while others facing similar circumstances drop out. We found that many related factors affect girls' ability to stay in school. Trends impacting whole communities can combine with household and individual factors to produce very different outcomes for girls from similar backgrounds. Dropping out of school is a lengthy process, and girls are often in a race against time to finish their education before social pressures such as pregnancy and marriage pushes them out of school. At the same time, many girls' time in school is interrupted by periods out of school because of different types of shocks and events that affect their ability to pay school fees.

Only very few projects commented on sustainability in their baseline reports and built sustainability firmly into their project designs. As mentioned above, most projects did not conduct a systematic political economy analysis at the design stage. As a result, they did not fully appraise the best entry points for complementing existing government initiatives or for creating strategic partnerships that could help sustain some of the interventions in the longer term.

Sustainability was prioritised at the start as a key logframe outcome that projects were required to deliver. However, in practice many projects started developing and implementing their sustainability strategies later on in the implementation process. At midline, projects are now facing a race against time to get enough stakeholder buy-in and investment, or develop enough resilience and capacity among partners to continue delivering the right activities and in the right ways to sustain their benefits.

The “holistic approach” taken by projects makes this even more difficult. Projects are hesitant or unable to identify individual interventions that have been particularly important and effective, and which as a result should be sustained. Many projects have argued that their interventions work best if delivered as the “holistic package” that was provided through the GEC funding. It is, however, highly unlikely that government stakeholders will have the resources, capacity, and political will to sustain complex intervention packages that involve various delivery partners and cut across different domains.

Equally, projects will struggle to convince government and other stakeholders to take over only a few of their interventions, unless there is clear and robust evidence of which interventions work best, for whom, how, where and with what effects. Furthermore, all stakeholders need to be convinced that in some form the benefits will outweigh the cost of their investment of time, effort and resources in a new approach. Projects need sufficiently granular evidence about the effectiveness of new approaches to meaningfully and credibly demonstrate their cost-effectiveness to government stakeholders in particular who face competing priorities for the scarce resources available to them.

Rigorous experimental and quasi-experimental evaluations require expert resources and stable conditions that are not always available on the ground.

The GEC’s Evaluation Strategy requires a particularly rigorous approach to evaluation and data collection. At the GEC learning events held in different regional hubs throughout the summer of 2016 many projects emphasised the rich data and learning that the GEC had produced. At the same time, however, the GEC has demonstrated the limits to using experimental and quasi-experimental evaluation designs and longitudinal surveys in contexts marked by a lack of reliable data (in particular geographical data), insecurity and instability, natural disasters, and resulting migration.

As discussed in Section 2.5, many SCW projects struggled to obtain data in line with the quality standards set by the GEC Evaluation Strategy. On one hand, contextual challenges such as insecurity, instability and migration made it difficult or even dangerous to track cohorts and sample control groups. On the basis of do no harm principles, DFID and the FM required six projects to drop their control groups for safety and ethical reasons. Several projects also had such high attrition rates that their samples have taken on a cross-sectional rather than longitudinal character.

On the other hand, projects had to rely on external evaluators who often had little experience of implementing similarly complex evaluations. Poor handling of identifier and address lists, the miscoding of variables, and changes in external evaluators all created additional hurdles to the collection of reliable, longitudinal data. A key learning point is that the implementation of such complex evaluation designs requires strong technical M&E capacity on the ground, as well as close supervision, guidance and quality assurance from centrally provided M&E technical support.

At the start, projects were required to rigorously select control groups that were well-matched to their treatment groups. SCW projects were asked to provide lists of possible control schools, frequently before their external evaluators were fully contracted. These lists were often provided by project staff without specific experience in planning and conducting quasi-experimental designs. Treatment schools were often selected for convenience (e.g. because projects were already working in these areas) and control schools were matched based on a few available variables, such as average regional enrolment rates and socio-economic indicators. Our analysis suggests that many treatment and control sampling points did not have sufficient geographic distance from each other, causing spill-over effects, and methodological “contamination” of the control group with girls living in a control school’s catchment area but being enrolled in GEC intervention schools.

This has reduced the power of the EM's and projects' samples and makes it more difficult to pick up effects on outcomes and barriers, even if they actually exist. In future programmes, these challenges could be mitigated by selecting control groups more carefully, possibly based on a pilot study to assess whether treatment and control groups match with regards to key characteristics and to ensure they are geographically clearly separated from each other. In addition, future EM designs might want to track a cohort of girls that is known to participate in project interventions, rather than a general population sample, to ensure that the treatment group has actually received treatment. Several projects have taken this approach in the SCW by measuring learning with girls in targeted grades only, which may have contributed to project results differing from EM results.

Projects' external evaluators struggled to produce good quality qualitative analysis and evidence of the strengths and weaknesses of specific project interventions. It is therefore not clear which interventions were most critical and most effective in enabling girls to go to school, attend school and learn.

The window-level evaluation was designed to measure aggregate effects across projects. So, the EM had to rely on projects' evaluation reports and data to understand what has worked and what has not at the project level. However, the analysis provided by projects' external evaluators often did not make the most of the available data by failing to merge baseline and midline household datasets and assess changes in barriers and intermediary outcomes overtime. They also struggled to triangulate the quantitative and qualitative data in effective ways, and did not test design assumptions systematically. In addition, the EM would have needed more detailed information about each project's intervention to unpack what was working, why, for whom and under what types of conditions.

Future evaluation designs should take into account that window-level conclusions are very difficult to draw for multi-country programmes in which each project targets a distinct population and context with a distinctive theory of change. To produce sufficiently meaningful learning about what works in addressing key barriers in a specific context requires a clear understanding of intervention mechanisms and their specific theories of change at the project level. The scale and diversity of the SCW means that this cannot be evidenced at the programme level. To do this, good quality qualitative evidence about what works, why, for whom, where and with what specific effects on learning outcomes is needed at the project level.

However, projects' external evaluators struggled to produce good quality qualitative analysis and evidence of the strengths and difficulties of specific project interventions. It is therefore not clear which interventions were most critical and most effective in enabling girls to go to school, attend school and learn. In addition, project-level evaluations would have benefitted from a more systematic assessment of the theory of change, by explicitly testing assumptions and hypotheses, tracking whether intermediary outcomes have been achieved, and monitoring changes along various steps in their results chains and impact logic.

5 Recommendations

SCW projects' activities end in July 2017. DFID and the FM are currently in the process of commissioning the next phase of the GEC. All of the current GEC projects are eligible to submit proposals for the new GEC Transitions Window, which is one of two funding windows that will make up the next phase of the GEC. The other window is called 'Leave No Girl Behind', and focuses on those girls who are the most marginalised.

The following recommendations have been developed with a view to specifically informing the next phase of the GEC. These recommendations are aimed at DFID. However, the actions required to deliver them involve DFID, the FM, grantees and the EM.

1. **Project designs need to be more evidently focused on directly delivering the required results.** There is a lack of evidence of that projects are focusing on problems that are critical to improving girls' learning outcomes. . Projects need to better demonstrate and describe how each type of intervention contributes to delivering improvements in girls' learning in time to deliver their agreed targets.

   Key Actions: at the design stage, with the support of the FM, projects should be required to describe in more detail the critical path between activities, intermediate outcomes and learning outcomes. This should include a clear explanation of which barriers will be affected by which project intervention or combination of interventions at each step to explain how and why this will specifically lead to improvements in girls' learning.

2. **Context analysis should be an integral part of the design process and M&E, particularly when adapting designs that have been tried and tested from elsewhere.** Marginalisation is very context-specific. Contextual factors have a significant effect on projects' success. They need to be tested, tracked, evaluated and reported on as an integral part throughout the project lifecycle.

   Key Actions: before baseline research starts, key assumptions about the project's context should be rapidly appraised and evidenced. This should include:

   - *political economy analysis* of the education system at national, provincial and local levels; and
   - *contextual analysis* of relevant factors outside the control that are most likely to affect the success of the project covering political, environmental, social and economic factors in particular.

   The baseline findings from this analysis should form the basis for ongoing project M&E and performance reporting.

3. **Projects should better demonstrate that they are coordinating with other programmes and actors working in SCW project environments.** The factors marginalising girls from education are multi-faceted cutting across different sectors and different parts of the education system. The most effective way of addressing such a wide range of problems is to join up with other actors who are able to contribute to projects' goals of improving girls' learning outcomes.

   Key Actions: at the design stage, projects should be required to map their strategic landscapes to identify which other government or non-government actors are planning or currently delivering activities that could help address the barriers that their target girls and communities face. This should form the basis for designing in and demonstrating strategic collaboration between GEC projects and other relevant initiatives and programmes.

4. **Gender analysis and gender action planning should be at the heart of GEC project design, delivery, M&E and reporting processes.** It is critical to the success of GEC projects that they are able to identify and track the extent to which girls are disadvantaged from achieving education outcomes compared to boys and how.

   Key Actions: from the start all projects should systematically gather primary and secondary evidence of gender parity in enrolment, attendance, retention and learning outcomes across grades, school phases at national, regional and local levels. GEC problem analysis and project design processes and targeting strategies should be driven by clear evidence of girls being significantly disadvantaged compared to boys. Boys as well as girls should continue to be the focus of project M&E – this should include testing the

learning outcomes of both boys and girls as an integral part of the GEC Evaluation Strategy and projects' M&E frameworks.

5. **Sustainability should be a central focus for the next phase of the GEC and structured to deliver effective sustainability strategies from the start.** Projects need to plan from the start who they will engage with and why, and set out what actions are needed to facilitate stakeholder buy-in and commitment to sustaining activities critical to the sustaining the benefits for projects' target girls.

   Key Actions: sustainability strategies should include:

   - *A government engagement plan* – setting out plans for engaging government at a national policy level and administratively at regional /provincial and local levels. From the outset these plans should clearly set out the roles and responsibilities envisaged for government and form the basis for ongoing negotiation and collaboration.
   - *A community engagement plan* – setting out plans for engaging different parts of the community, its roles and responsibilities for continuing key activities after the end of the project. This should include an assessment of the key driver and barrier to continued community engagement and capacities that need to be developed throughout the life of the project.
   - *A stakeholder coordination plan (including government, DFID country offices and other donors)* – setting out plans for coordinating with other programmes and initiatives both within the education sector and across other sectors (e.g. WaSH, livelihoods, health, child protection etc.) relevant to girls' education.
   - *A commercial sustainability plan* – this should set out key input costs against each of the project's outputs /types of intervention. This should enable projects to measure their key output unit costs as a starting point to assessing and demonstrating to stakeholders the commercial viability of their sustainability plans.
   - *Match funding and leverage plans* – this should set out the project's plan for influencing stakeholders before and after the project has started. There should be a clear distinction between match funding and leverage:
     - *Match funding* should be defined as funding attracted *before* the start of the project from partners and stakeholders who have bought into the project concept and want to financially support the delivery process; and
     - *Leverage* should be defined as additional resources that are committed by stakeholders because of the project's activities and influence during delivery – this should act as a better indicator of stakeholder investments in sustaining activities beyond the end of the project.

6. **A counterfactual approach to impact evaluation adds value to programme learning and accountability, but DFID and the FM should reconsider how it is carried out and by whom.** Designing and delivering a counterfactual approach to impact evaluation has involved a large investment of time and resources by DFID, the FM, EM and projects. The process has faced significant challenges as a result of: a lack of technical capacity among projects, their external evaluators and local research organisations; adverse environmental conditions; and changes in project designs and targeting. Yet, impact evaluation at the project-level appears to be the most appropriate unit of analysis due to the variety of projects' contexts, strategies and intervention types. So, a potentially more effective and better VfM alternative could be for a single agency to carry out all project-level impact evaluations. While evaluations commissioned by projects focus on assessing the processes, performance and effectiveness of their interventions.

7. **A project M&E capacity building plan should be developed from the start.** This phase of the GEC has demonstrated that there are beneficial spill over effects from building the internal M&E capacity of projects.

   Key Actions: from the start, the FM conduct a rapid needs assessment with regards to projects' internal capacity to support the scale and level of research, monitoring and evaluation required – this should form the basis of a comprehensive capacity building plan to strengthen projects' ability to specify, commission, manage and conduct M&E.

8. **Projects' M&E frameworks need to deliver robust evidence of which types of interventions are most effective in delivering girls' learning outcomes.** Without robust evidence of what works best, how and why, projects will find in difficult to improve and correct their project designs and deliver processes. Projects will also struggle to get buy-in from stakeholders to sustain key activities.

   Key Actions: The following changes to M&E frameworks should help projects better measure what works:

   - introducing *intermediate outcome targets* would enable the FM and projects to track and measure the contribution of different types of interventions towards improvements in girls' learning outcomes;
   - a stronger focus on *qualitative research and analysis* will projects identify the relative effectiveness of different types of interventions; and
   - a greater emphasis on *process evaluation* would help projects distinguish between design or implementation success /failure when evaluating their effectiveness and impact.