coffey >
A TETRA TECH COMPANY

Endline Evaluation Report Strategic Partnerships Window

Final Report
Evaluation Manager Girls' Education Challenge Fund
October 2017

![coffey - A TETRA TECH COMPANY]

Endline Evaluation Report

Strategic Partnerships Window

Final Report

Evaluation Manager Girls' Education Challenge Fund

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:

Abbreviations and Acronyms

ASER	Annual Status of Education Report
CBN	Central Bank of Nigeria
DAC	Development Assistance Committee
DFID	Department for International Development
DHS	Demographic and Health Surveys
DP	Discovery Project
EFA	Education For All
EGMA	Early Grade Math Assessment
EGRA	Early Grade Reading Assessment
EM	Evaluation Manager
ENGINE	Educating Nigerian Girls in New Enterprises
FCT	Federal Capital Territory
FGM	Female Genital Mutilation
FM	Fund Manager
GEC	Girls' Education Challenge
GNI	Gross National Income
GPE	Global Partnership for Education
HIV/AIDS	Human Immunodeficiency Virus / Acquired Immunodeficiency Syndrome
ICT	Information and Communications Technology
ISG	In-School Girls
IW	Innovation Window
KCPE	Kenya Certificate of Primary Education
KES	Kenyan Shilling
LC	Learning Centre
LSC	Learning Space Coordinators
M&E	Monitoring and Evaluation
ODA	Overseas Development Aid
ORB	Opinion Research Business

OSG	Out-of-School Girls
PE	Performance Evaluation
PTA	Parent-Teacher Association
PwC	PricewaterhouseCoopers
Q-E	Quasi-Experimental Method
RCT	Randomised Controlled Trial
RTI	RTI International
SBMC	Sensitisation and Mobilisation of Community Members
SCW	Step Change Window
SD	Standard Deviation
SDG	Sustainable Development Goals
SPW	Strategic Partnerships Window
SRH	Sexual and Reproductive Health
ToC	Theory of Change
TSC	Teachers' Service Commission
TV	Television
UN	United Nations
UNESCO	United Nations Educational, Scientific and Cultural Organisation
UNGEI	United Nations Girls' Education Initiative
US	United States
USAID	United States Agency for International Development
WASH	Water, Sanitation and Hygiene

Country and Locations Abbreviations

The following abbreviations are used for countries in tables in this report:

Ken	Kenya
Nai	Nairobi
Waj	Wajir
Gha	Ghana
Nig	Nigeria

Contents

Executive Summary

Introduction to the GEC

The Department for International Development (DFID) launched the £355 million Girls' Education Challenge Fund (GEC) in 2012, setting an ambitious target of reaching one million marginalised girls[1] by March 2017[2].

The Business Case[3] for the GEC highlighted that while there was a lot of evidence around *"(1) the benefits of female education, (2) barriers to participation and (3) successful interventions"*, less was known about which barriers were marginalising girls from education, and questions remained over how best to intervene at scale. This was the underlying rationale for a fund that, not only funded interventions, but challenged organisations to identify the root causes of educational marginalisation and rigorously collect evidence on their projects to ascertain what worked well, why it works, for whom and under what circumstances.

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

The GEC consisted of three funding windows: (1) the Step Change Window (SCW), where projects were awarded up to £30 million each to deliver approaches to improve girls' education at scale.; (2) the Innovation Window (IW) where projects were awarded up to £2 million each to test and pilot new approaches; and (3) the Strategic Partnerships Window (SPW).

This report presents the findings from the end-line evaluation of the Strategic Partnerships Window.

The Strategic Partnerships Window

The aim of the SPW was to build partnerships with the private sector that combined social and business approaches to support girls' education in a sustainable way. The GEC Business Case (2012) had the premise that "*there is potential for the non-state sector to play a more significant role in reaching marginalised girls. But there is a lack of evidence on how best donors can scale-up support and work in a sustainable way with the full range of non-state providers. There is a need to close the knowledge gap, to lead to more effective relationships with non-state education providers. The GEC could play a role in doing this.*"

Private sector involvement, and the resources leveraged as result of these partnerships was one of the critical success criteria for the GEC. Seeking matched funding from private sector partners was considered an important avenue to try to address the *significant funding gap for basic education* – UNESCO estimates[4] that an annual finance gap of US$39 billion in low income countries needs to be filled to achieve the SDG Education 2030 targets. DFID considered that public-private partnerships could help "*to embed the focus on girls' education in other parts of the international system, helping to ensure sustainability of support."*

The SPW has a distinct focus on developing partnerships where there is an overlap between social and commercial aims. The window is comprised of four partnerships: Coca-Cola; Discovery Communications; Avanti Communications and Ericsson. These partners implemented projects in Kenya, Ghana, Nigeria and Myanmar, with a total budget of between £7 million and £27 million. DFID provided approximately half of the budget, with partners providing the other half as match funding.

Purpose of the SPW Endline Evaluation Report

While the GEC launched in May 2012, the funding windows were staggered, starting with the Step Change Window (May 2012), the Innovation Window (July 2012) and lastly the Strategic Partnerships Window (in October 2013).

---

[1] Girls' Education Challenge: http://devtracker.dfid.gov.uk/projects/GB-1-202372
[2] Ibid
[3] DFID (2012), Girls' Education Challenge, Business Case Version 4; London: DFID.
[4] UNESCO (2016) 'Global Education Monitoring Report – Education for people and planet: creating sustainable futures for all'

The first window saw the announcement of partnerships with Discovery Communications and Coca-Cola. A second phase was launched in July 2014 with further partnerships with Avanti Communications and Ericsson.

This endline evaluation report follows the SPW Baseline Report from January 2016[5].

This report includes data and findings covering three SPW projects: Coca-Cola; Avanti Communications; and Discovery Communications. The fourth SPW project, Ericsson's project in Myanmar, is excluded from the report as it had not finalised its endline report and data analysis in time to be included in this evaluation. Ericsson was also not included in the SPW Baseline Report for the same reason.

This evaluation focuses on evaluating SPW projects' effectiveness and impact on marginalised girls' education outcomes. It does not test the assumptions underpinning the rationale for the SPW set out in DFID's Business Case. The findings, conclusions and recommendations in this report reflect the unique characteristics of these private sector partnerships. Further lessons about the process of setting up and delivering these partnerships in the SPW are covered in the EM's Process Review[6] and in the FM's Strategic Partnerships Consultation[7].

The FM's Strategic Partnership Consultation Report provides valuable insights into what worked well (and less well) with regards to the process of developing the partnerships in light of the original ambition. For the purpose of context, the report usefully summarises DFID's ambitions as follows:

> *"The Strategic Partnerships Window was intended to work with new partners, in new ways, in multiple countries, with exacting M&E standards and with limited time for design. The level of ambition has meant that challenges were inevitable, but good progress has been made. The balance of the 'partnership' versus the 'project' benefits continues to shape decision-making, engagement and support to the Lead Partners."*

It is important to recognise the context in which these partnerships were built. These types of public-private partnerships were in themselves new and untested in education and development. Private sector partners were expected to come with solutions that could be both commercially sustainable and deliver challenging education outcomes for marginalised girls. The purpose of this endline evaluation of the SPW is to produce reliable evidence of the programme's effectiveness and impacts on the marginalised girls targeted by the three projects. As such, it does not cover the public-private partnership building process that inevitably shapes the type of project designs that were implemented. However, the context in which these projects were designed and delivered should be taken into account when reviewing the findings, conclusions and recommendations arising from this evaluation.

It is anticipated that DFID, the FM and SPW partners will use the findings and lessons learned from this evaluation to inform the further development of these projects, as three of the partnerships (Coca-Cola; Avanti Communications; and Discovery Communications) will continue to receive DFID funding through the new GEC Transitions programme (GEC-T). It is hoped that this report gives transferable findings for a wider audience including donor agencies, national governments and other policy-makers, if they seek to engage private sector partners in improving marginalised girls' education outcomes.

Evaluation approach

We assess the overall impact of the projects on the targeted girls in 'treatment' communities and schools. The SPW evaluation design relies solely on evidence collected and reported by the SPW projects. We reviewed and undertook a meta-analysis of the reports, datasets and outcome spreadsheets submitted by SPW projects. The process adopted aims to synthesise the evidence provided by projects to report on evaluation findings at the window level across the three projects. In addition, we drew on relevant secondary data and existing literature to help triangulate our analysis or to allow us to explore our findings.

---

[5] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/494679/Baseline-Strategic-Partnerships-Window-13January2016.pdf
[6] Coffey (2015), GEC Process Review Report
[7] PwC (2015) Strategic Partnerships Consultation

Key findings

Reach

The SPW projects aimed to reach not just girls, but marginalised girls. The projects' definitions of marginalisation did not change between baseline and endline and included aspects of socio-economic, educational and geographic marginalisation. While projects demonstrated an increased understanding of marginalisation in their project areas in their endline reports, there was no reported evidence that projects adapted their targeting strategies or activities to reflect these changes.

According to the projects' endline reports and the FM's programme management information, SPW projects reached 380,350 girls and 426,920 boys. This includes boys that were reached despite not being specifically targeted, and therefore were unlikely to have received the same intensity of intervention as the targeted girls.

Discovery considered all the girls in their project areas to be marginalised, and delivered activities based on a whole school approach. Avanti, in contrast, mostly adopted a 'whole school' approach yet did not consider all girls to be marginalised, and used a stipend component to specifically target the most marginalised girls. Avanti estimated that 10,497 families and 115 merchants have been reached. To some extent, the whole-school approach chosen by Avanti and Discovery explains the large number of children reached, and in particular the large number of boys participating in SPW activities.

Avanti's stipend component used a combination of targeting criteria based around school attendance and validation by schools to target those perceived to be most at risk of dropping out of school. The project reports that 55% of stipend recipients improved their attendance, but there remained unmet financial barriers to school attendance and factors such as 'sickness', which accounted for a large amount of girls missing school.

Coca-Cola estimated that they reached about 40,000 community members, but their endline evaluation report provided a limited amount of information about which groups of girls make up the most marginalised in the project, or if they reached the most marginalised girls in their regions and communities. However, Coca-Cola was the only project that targeted specific girls for all of its project activities. Girls undertook self-screening to participate in the project based criteria including: being an orphan; a parent being sick; having a disability; or being married before the age of 18. The EM reanalysis of their data suggests that Coca-Cola had some success in reaching the most marginalised, particularly orphans /girls in single-headed households.

Impact on learning

Impact on learning was assessed by comparing the change in learning outcomes from baseline to endline between the projects' treatment and control groups. This difference was not statistically significant for most learning outcomes in most project areas. There is little evidence that SPW projects had an impact on learning outcomes.

The only positive, and statistically significant impact on literacy was found in the Discovery project in Nigeria, though conversely they had a statistically significant negative effect on girls' numeracy outcomes in Kenya and Ghana. The positive effect in Nigeria was found for girls whose literacy skills were initially very low, while the negative effects came in areas where numeracy skills were initially relatively better.

Avanti did not have a statistically significant impact on learning outcomes, but the full package of activities they delivered had a greater effect than the various combinations that excluded some elements. Tutoring appeared to improve literacy for high achieving girls and improve numeracy for low achieving girls. Stipends appeared to have the opposite effect.

Coca-Cola had no effect on in-school girls learning in terms of their literacy or numeracy. Coca-Cola's out-of-school girls represent more than half of the project's beneficiaries, yet their literacy and numeracy of these girls was not tested because the project's theory of change had a greater focus on improving their financial literacy, which showed some improvements.

Impact on attendance

Across the SPW, there were issues with the quality of the attendance data, which limit our ability to generate conclusive findings in comparison to control groups. At endline, Avanti's attendance results were inconclusive. School attendance records for control schools were not available or not properly kept.

For Coca-Cola, the attendance data provided was collected at their Safe Spaces rather than in schools – they were calculated based on an average of attendance rates during the nine-month period for each cycle of the course.

Discovery achieved its attendance target at endline in Nairobi (Kenya) and in Nigeria, whereas Coca-Cola nearly achieved it for their Safe Spaces for both of the cycles. In addition, Coca-Cola reported encouraging results for their out-of-school girls, most of whom graduated from both of the nine-month training cycles.

Effectiveness of SPW projects' activities

This section discusses the effectiveness of the SPW's activities in addressing the most important barriers that girls faced in accessing a quality education.

School-related factors

All three projects designed and implemented interventions addressing school-related barriers. This mainly consisted of the provision of ICT equipment to schools (Discovery and Avanti) and teacher training (all projects). It is found that school-related factors are still prevalent barriers to education at endline.

Projects found that perceptions among caregivers of the quality of teaching (Discovery), and perceptions of having adequate teaching materials and learning aids have improved (Discovery and Avanti). Projects also found that ICT equipment created a more interesting environment for teaching and learning, and improved teachers' confidence and the teaching and learning process. However, both Discovery and Avanti found that teachers and schools faced serious challenges, which limited the usefulness of the equipment and the learning content, and limited their effects on the education process and outcomes.

Our reanalysis of the project data found evidence that teaching practices did not necessarily improve across countries and across practice areas. Where they did improve, the evidence suggests that some teaching approaches were more suitable and therefore better in certain contexts compared to others.

Poverty-related factors

Poverty-related factors are still prevalent barriers to education at endline. Only Avanti designed and implemented a core activity (i.e. the provision of stipends) to directly address poverty. Coca-Cola focused on income-generating activities, and facilitated saving groups and other economic and financial-related interventions, which were not intended to directly affect the barriers to an academic education that their target girls faced.

Projects reported that direct interventions to offset the cost of schooling increased attendance, while projects' indirect interventions (e.g. income-generating activities) did not seem to affect attendance or girls' education outcomes. Projects also reported that while there was some evidence that caregivers might have changed their attitudes around the benefits of girls' education, this did not translate into an actual decrease in girls' housework commitment. This meant that it did not translate into improved attendance and enrolment in school.

Our reanalysis of project data suggests that poverty remains one of the most important barriers to girls' educational achievement and that the projects had little effect in this area.

Girls' aspirations and decision-making

There is some evidence that girls' aspirations and interest in schools has improved from baseline to endline. Girls' clubs and safe spaces in particular appear to have improved girls' wellbeing and confidence, but girls still face challenges relating to decision-making powers around schooling and marriage.

Attitudinal factors

There is some evidence that Discovery managed to increase parents' engagement in children's education. However, there is mixed evidence suggesting that community outreach and awareness raising activities carried out by Discovery and Coca-Cola improved attitudes towards girls' education. The evidence also suggests that parents' engagement was limited to those parents who were already interested in education, 'better-off' households, and those who were qualified or educated to a certain level.

Our reanalysis of projects' data found that approaches to changing attitudes that were not targeted at particular individuals or specific issues were not successful. Any meaningful effort to change attitudes needs to identify a specific problem relating to attitudes to have an effect on behaviour beyond improving parental awareness of the benefits of girls' education.

Violence-related barriers

At baseline, both Discovery and Avanti found evidence of violence-related barriers to education that were prevalent in their target communities. Only Discovery aimed to indirectly address some of these factors through gender-responsive training for teachers and through community workshops.

There is no evidence though that these activities have been effective in improving education outcomes. Discovery provided evidence that violence, particularly corporal punishment and safety on the way to school, has not improved.

Sustainability

All projects reported that they engaged and mobilised communities to support the delivery of their activities. While there is evidence of community support and some commitment to continue projects' activities, there is also strong evidence that communities do not have the financial resources to maintain and continue projects' activities in their current form. Avanti and Discovery in particular rely on schools to provide the space and teachers needed to continue using various learning resources that have been introduced and implemented. However, sustainability strategies that rely on teachers and schools are undermined by high teacher mobility and a lack of financial support from communities and Ministries of Education.

All SPW projects have engaged government bodies at national, regional and district levels with some success in terms of aligning with policy and influencing policy change. However, this engagement has not yet translated into concrete plans to fund projects' activities so that they can continue. This requires a lot more time and continued advocacy for institutional change. All three SPW projects will continue to receive support (in one form or another) through the successor programme to the GEC, the GEC Transitions (GEC-T) programme. Without this support it is highly unlikely that many SPW activities would have continued as they stand now.

Key conclusions

SPW projects had little or no impact on literacy and numeracy outcomes of the marginalised girls that they reached. A combination of weaknesses in projects' designs and the ways in which they were implemented contributed to this finding. In addition, contextual factors, such as the Ebola outbreak in Nigeria, affected implementation.

The evaluation of each SPW project's impact on learning outcomes was evaluated through a quasi-experimental design, even though all of the projects' endline evaluation reports refer to randomised control trials. The main finding from this evaluation was that the literacy and numeracy results of projects' beneficiaries did not improve more than the results of comparable girls who were not beneficiaries.

Across the SPW projects – disaggregating Discovery's project into the three project areas where it worked – the only project area where there was a positive, statistically significant effect on learning was in the literacy scores of girls supported by Discovery in Nigeria, but this effect was small. In the two other project areas supported by Discovery, in Nairobi and Ghana, there was a statistically significant negative impact on numeracy.

Avanti had no effect on literacy and numeracy. Coca-Cola had no effect on literacy or numeracy with some results proving inconclusive due to a lack of statistical tests being undertaken.

It is worth exploring the causes for the small statistically significant impact in Nigeria's reading scores and for Discovery's negative results elsewhere. The girls in Discovery Nigeria's project areas had by far the lowest reading levels at baseline, suggesting Discovery's approach can help improve literacy, but only for students at the bottom of the learning spectrum. This resonates with the EM's findings from the midline evaluations of the SCW and IW that showed these portfolios as a whole had the greatest impact on those girls with the lowest levels of literacy.

Projects focussed on marginalised regions, not always on the most marginalised girls. The evidence suggests that where interventions targeted more marginalised girls, activities have been more effective.

Projects applied different targeting approaches. While Discovery and Avanti targeted whole schools, Coca-Cola targeted specific communities. In the case of Discovery and Avanti, the schools were identified based on marginalised communities. Avanti made a distinction between "educational marginalisation" and "poverty marginalisation" in their M&E plans and used poverty as a primary factor to identify marginalised areas in Kenya. Furthermore, despite trying and ensure a high level of marginalisation in the project's schools, Avanti's sampling frame of schools included only those that had access to electricity and participated in Kenya's sanitary towel

programme[8]. Restricting selection to those schools with access to electricity may have excluded some of the poorer areas from the sampling frame. Therefore, it is hard to conclude Avanti targeted the most educationally marginalised communities.

In cases where communities were educationally and economically marginalised, we found that some improvements in education outcomes have been achieved. Additionally, in the cases where interventions were more targeted, such as Coca-Cola's project, or the stipends component in Avanti's project, there is evidence that marginalised girls have been reached, and/ or that some results have also been achieved as a consequence.

At the window level, SPW projects have reached more boys than girls. However, there is no evidence that there was an increase in gender disparities as a result.

Across the SPW, projects reached more boys (426,920) than girls (380,350). The underlying assumption behind the GEC is that girls face gender-specific challenges, which put them at a disadvantage educationally compared to boys. Although the GEC Business Case states that the programme is designed to address disparities in girls' education compared to boys, it was not set up at the start to explicitly measure how or to what extent projects affected gender disparities in girls' education outcomes.

Discovery's and Avanti's core activities predominantly involved a whole school approach using education technology. The ways in which these activities were delivered meant boys could not be excluded in the process of delivering support to their target girls. In this sense boys were reached and supported at no additional cost.

The original Theory of Change for the GEC in the Business Case defined how the programme would be delivered rather than the changes that would be experienced by girls as a result. In this context, no gender studies have been conducted by any GEC projects at endline, including the SPW projects, and so cannot provide any conclusions about changes in gender disparities in education outcomes.

Projects did not always focus the design of their interventions to achieve the specific educational outcomes targeted by the GEC.

Discovery's and Avanti's main project activities involved the introduction of the ICT equipment and improving pedagogy in schools. Coca-Cola provided business and finance training to out-of-school girls and additional instruction to in-school girls after school. These approaches responded to the expectation laid out in the GEC Business Case that private sector investment would *explore new ways of supporting girls in education*. However, the evidence suggests that these approaches were generally not sufficiently focused on improving school attendance and the learning outcomes (i.e. literacy and numeracy) as intended by the GEC.

The largest group of beneficiaries in Coca-Cola's intervention was out-of-school girls. These girls did not receive any significant literacy or numeracy training. Instead, this intervention was focused on improving their financial literacy with further opportunities for potential employment in Coca-Cola's value chain. Girls who were enrolled in school and were beneficiaries of Coca-Cola's intervention only received two hours of tutoring per week to cover all the course content included in Coca-Cola's curriculum – this included business skills and finance, leaving less than one hour to cover academic content including numeracy and literacy. From a programme design standpoint, this approach placed too little emphasis specifically on the GEC's key learning outcomes.

The core elements of Discovery's programme focused on improving pedagogy – introducing ICT equipment to create a more engaging classroom, and training teachers in student-centred and gender-responsive teaching methods. However, neither the content introduced through ICT nor the new teaching methods focused on methods to improve specific skills in literacy or numeracy. Discovery's learning content, conveyed through the educational videos, mostly focused on science and social studies rather than on literacy and numeracy. Teaching methods focused on student engagement and fostering creativity. While it is possible that improving the overall quality of teaching could improve specific outcomes such as literacy and numeracy, the project design, content or teaching methods did not focus on these outcomes. The lack of statistically significant impacts on literacy and numeracy at endline supports our conclusion that pedagogical interventions that are not specifically designed to improve literacy or numeracy are unlikely to have that effect.

---

[8] Avanti's M&E framework.

Avanti was the only project in the SPW that focused specifically on improving numeracy in their project design. Since we also found no impact of this project on learning outcomes, this is most likely explained by other factors relating to its implementation and context.

Projects did not sufficiently adapt their projects to the local context with respect to their design or implementation strategies.

Projects generally appear to have rolled out a blanket approach to designing interventions in various locations and contexts. There were some exceptions – such as Discovery's flexible approach to running the girls' clubs, or projects' community mobilisation and parents' engagement in education activities, which were generally designed from the bottom-up.

The core activities tended to be solution-driven, involving particular types of equipment, products and solutions that projects had already developed and used elsewhere. The SPW challenged private sector partners to respond to the GEC with potential commercially viable solutions. However, the evidence suggests that these solutions were not sufficiently adapted to suit the context, needs and priorities of the girls, schools and communities that projects were working with to have a significant effect on literacy and numeracy.

For example, it is not clear how Discovery considered differences in contexts when rolling out its activities in its three target countries. For example, Ghana appears to be a special case when it comes to parents' engagement and perceptions about girls' education in comparison to the other countries and contexts where Discovery worked. The educational levels of schools were dramatically different across countries as well. At baseline, a large majority of students in schools in Nigeria could not read a single word, whereas students in Kenya were mostly able to read, but were somewhat behind international benchmarks. Moreover, the cultural context of the different areas varied greatly. Given these differences, it would be have been surprising if the same package of pedagogical techniques and learning materials would respond to the needs of the students in each area.

Problems in implementation of projects may also have reduced their effectiveness.

The reports show numerous challenges that projects faced during implementation, many of which persisted until the end of the project. For example, Avanti's approach involved substituting the teacher with an individualised learning platform. The project's evidence suggests that the approach did not get sufficient buy-in and engagement from teachers, so the technology was underused by students.

There is also evidence that the number of hours per day in which the ICT equipment was functional decreased to an average of 2.6 hours per day (for a school of about 1000 student with only 25 computers). Some of the underlying causes for this included electricity bills increasing and schools not being able to pay them, or technical problems. The ICT equipment Discovery installed faced similar problems. Students were not always able to see the televisions, and the content was not always appropriate for what the teacher wanted to accomplish.

One clear finding from the SPW is that ICT equipment can only act as a tool to support teaching, and cannot effectively substitute good teaching practice. Ensuring that ICT equipment is effective appears to be a more intensive task than anticipated, requiring monitoring, maintenance, and mechanisms to ensure its fitness for purpose.

Projects are unlikely to continue without further DFID support through the successor GEC-Transitions programme, and this depends on the design and reach of the new projects.

While sustainability remains a challenge, the rationale for the SPW was driven by the need to attract commercially sustainable private sector investment in girls' education. This succeeded, but the investment has not delivered sufficiently on educational outcomes, sustainable activities or benefits. All the projects effectively engaged government bodies at multiple levels and also engaged and mobilised schools and communities to the designs that they wanted to implement. However, at endline there are no clear plans among communities, schools or government ministries to sustain continued investment in the activities specifically delivered by the projects. The sustainability strategies currently put forward by projects are overly reliant on additional school and community resources that are fragile at best.

However, all three SPW projects reviewed in this report will continue to receive support (in one form or another) through the successor programme to the GEC, the GEC Transitions (GEC-T) programme. The extent to which girls, schools and communities will continue to benefit from GEC-T will depend on the subsequent design of the SPW projects which are now tasked with continuing to support the same beneficiaries supported in the GEC. There is a risk that resources will be reallocated to support older age groups to ensure continued benefit for existing

beneficiaries, while previous GEC-funded activities supporting younger age groups or lower grades may not be sustained as a result of a change in focus of the projects' designs. It is plausible that with more time these projects could increase in effectiveness, but the evidence strongly suggests that significant changes need to be made to their designs and delivery processes, and more time in itself would not be sufficient. Only activities that are effective should be sustained, and GEC-T provides the SPW projects with an opportunity to re-focus their designs and correct their delivery processes to have a greater impact on outcomes for marginalised girls.

There were significant issues with the quality of quantitative and qualitative data, analysis, and reporting across the SPW. A more prescriptive approach is needed to improve consistency and comparability.

The type and quality of data collected and submitted by each project varied greatly. Furthermore, the quality of quantitative and qualitative analysis reported in partners' endline evaluation reports was poor overall. The lack of sufficient quality and consistency in the evidence that partners submitted made it difficult to fully understand, compare and generalise about how and why different types of interventions worked, for whom, in what contexts and with what effects. If external evaluators are conducting project evaluations in GEC-T then it seems clear that a more prescriptive approach is needed with greater technical oversight to deliver improvements in data quality, analysis, reporting, learning, consistency and comparability across the programme.

In the relatively short time that the SPW has been delivering its activities it has had little evident impact on the education outcomes of the marginalised girls that projects targeted. From a 'project' perspective, the designs and delivery processes need to change to have a greater impact, specifically on girls staying in school and improving their literacy and numeracy. In the context of developing new and innovative public and private sector 'partnerships' that focus on girls' education outcomes, DFID, the FM and the partners have progressed and completed the first phase of the GEC. However, the evidence from this endline evaluation suggests that these partnerships need to be rebalanced to place greater weight on the education outcomes of the marginalised girls that DFID is continuing to support through GEC-T.

1 Introduction

1.1 Background to the GEC Strategic Partnerships Window

1.1.1 Context and rationale

Changes in global conditions affecting the education of marginalised girls

Every child has the right to 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 with girls who have only completed primary school.[9] 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 USD 10, and around one-third of the decrease in adult mortality since 1970 comes from improvements in the education of girls and young women.[10]

In launching the Girls' Education Challenge (GEC) Fund in 2012, DFID signalled its commitment towards 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[11] and 3[12], which specifically relate to universal primary education and gender equality. By the time the final Education for All (EFA) Global Monitoring Report (2000-2015) was published, more girls were in school, staying there longer, and learning more whilst attending[13].

The last decade has seen global improvement in universal primary education coverage with a majority of children, both boys and girls, entering primary education in most countries around the world. Although primary school enrolments for girls have improved along with boys, school completion rates remain equally low for both boys and girls. Around 264 million children and young people were estimated as being out of school by the end of 2015 – this includes 61 million children of primary school age (6-11 years), 62 million adolescent children of lower secondary school age (12-14 years) and 141 million children of upper secondary school age (15-17 years)[14]. 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 primary school level.

Improved access to education has not resulted in improvements in learning for many children. Levels of learning remain very low, even for those enrolled in school[15] - 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 consistency, inclusivity and quality of education worldwide[16].

In 2015 it was estimated[17] that children in conflict affected countries, who make up 20 percent of the world's primary school aged children, account for around 50 percent of primary aged children who are out of school and 36 percent of all children not in schooling[18]. In fragile and conflict-affected states and in crisis situations the provision

---

[9] 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.

[10] 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.

[11] 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.

[12] 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.

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

[14] UIS and GEM (2017), 'Reducing global poverty through universal primary and secondary education? Policy paper 32 /Fact Sheet 44. Paris: Global Education Monitoring Report'.

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

[16] 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.

[17] UNESCO (2015) 'Fixing the broken promise of Education for All. Findings from the Global Initiative on Out-of-School Children.' UNESCO.

[18] UNESCO (2015) 'Education for All 2000-2015, Achievements and Challenges.' UNESCO

of basic education services becomes difficult or impossible. Sub-Saharan Africa is the region with the highest out-of-school rates for all age groups, with three regions home to nine of out ten out-of-school adolescents: Sub-Saharan Africa (26 million) Central Asia and Southern Asia (20 million) and South-east Asia (8.5 million).[19]

Persistent under-investment in education

A key rationale[20] 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 2017, under-investment in education persists and ODA to education is yet to return to 2010 levels. 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 percent in 2013, aid to basic education declined by 7 percent,[21] and the overall volume of aid to education fell by about US$600 million between 2013 and 2014[22]. The UN Educational Scientific and Cultural Organisation (UNESCO) says ODA for education is now 4% less than in 2010. UNESCO also reports[23] that:

- there will be an annual finance gap of US$39 billion in low income countries to achieve the SDG Education 2030 targets; and

- the education finance gap can be filled if select donors give 0.7% of GNI to aid, and 10% of that to basic and secondary education.

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 percent), which prevents those children who are most marginalised from accessing a quality education.

UNESCO also highlighted the need for civil society and the private sector to play important roles in financing, implementing and ensuring mutual accountability to achieve the education targets set out in the SDGs. This particularly resonates with DFID's rationale for establishing strategic partnerships through the GEC. The need for diverse funding sources for education is also emphasised in the Education Commission's report[24], which calls for increased funding for education including ODA, emerging donors, non-concessional loans and funding from the private sector. The commission estimates that total financing for education needs to rise by around 11 percent a year to US$89 billion by 2030.

Changes in the global policy response to education

The GEC SPW Baseline Report was published in January 2016[25], a year after 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 girls. However, a greater focus is needed on the quality of education provided in order to achieve effective learning outcomes. 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, 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 is the first global fund to prioritise education in humanitarian settings. In 2016 the Education Commission also proposed an International Finance Facility for Education (IFFEd)[26] that could mobilise an additional US$10 billion annually for

---

[19] UIS and GEM (2017), 'Reducing global poverty through universal primary and secondary education? Policy paper 32 /Fact Sheet 44. Paris: Global Education Monitoring Report'.

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

[21] 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.

[22] UNESCO (2016) 'Global Education Monitoring Report – Education for people and planet: creating sustainable futures for all'

[23] UNESCO (2016) 'Global Education Monitoring Report – Education for people and planet: creating sustainable futures for all'

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

[25] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/494679/Baseline-Strategic-Partnerships-Window-13January2016.pdf

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

education by 2020. The model brings together public and private donors alongside international financial institutions such as the World Bank and regional development banks to create low-interest finance packages for lower-middle-income countries, linked to countries increasing their own level of investment in education and carrying out education sector reforms.

The Girls' Education Challenge Fund and its extension

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

- the Step Change Window (SCW);
- the Innovation Window (IW); and
- the Strategic Partnerships Window (SPW).

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

The current GEC programme ended in March 2017, with individual projects' contracts ending at various points between early 2017 and March 2017. At the Girls' Education Forum in London in July 2016, DFID announced it would provide a further £100 million of funding to the GEC[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. This support will further 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 relevant life and work skills 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 – e.g. 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 ToC and DFID's Business Case for the GEC assumes that there are educational barriers that affect boys and girls, but that girls face a number of additional, gender-specific challenges which put them at a disadvantage in comparison with boys. Although the GEC Business Case states that the programme is designed to address disparities in girls' education compared to boys, it was not set up at the start to explicitly measure how or to what extent projects affected gender disparities in girls' education outcomes. The original Theory of Change for the GEC in the Business Case defined how the programme would be delivered rather than the changes that would be experienced by girls as a result of the programme's interventions. To inform the GEC Evaluation Strategy the EM developed a ToC that more explicitly shows the typical links between different types of barriers and the GEC's outcomes, 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 lack of appropriate school facilities and distance to school; and institutional or political barriers such as lack of equity in public service provision, or a lack of influence of and accountability to marginalised groups.

---

[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'.

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 Overview of the GEC Strategic Partnership Window

This report will focus on the SPW. Its aim is to build partnerships with the private sector that combine social and business approaches to support girls' education in a sustainable way. The SPW aims to test new approaches and generate evidence about public-private partnerships, in addition to improving the lives of marginalised girls through education. The GEC Business Case (2012) expected the SPW to leverage private sector involvement and resources to explore new ways of supporting girls in education. Private sector leverage was identified as one of the seven critical success criteria. Match funding from private sector partners was considered an important strategy for addressing significant funding gaps for basic education that were described in the previous section. Public-private partnerships were also considered a particularly effective way of ensuring sustainability by targeting new investment in the non-state sector.

The total budget of the SPW projects is about £60 million; roughly half of this comes from the private sector partner and half is provided by DFID. Together, the SPW projects aim to support 130,000 marginalised girls across four DFID priority countries: Ghana; Kenya; Nigeria; and Myanmar.

While DFID officially launched the GEC in May 2012, the launch of each funding window was staggered, starting with the Step Change Window in May 2012, followed by the Innovation Window in July 2012 and lastly the Strategic Partnerships Window.

The first phase of the SPW began in October 2013 with the announcement of a partnership with Discovery Communications and Coca-Cola. The second phase launched in July 2014 with the announcement of a partnership with Avanti Communications and Ericsson. All GEC projects, including the strategic partnerships ended in March 2017. All funding windows have taken longer than expected to contract partners or organisations (and consortium leads), who in turn encountered delays in completing their M&E frameworks and baseline research.

In comparison to the other GEC windows, the SPW has a distinct focus on developing partnerships with the private sector to find opportunities where there is an overlap between the social and commercial business cases to address barriers to girls' education. The window is comprised of four partnerships implementing projects in Kenya, Ghana, Nigeria and Myanmar, each with a budget of between about £7 million and £27 million.

- **Coca-Cola Company's ENGINE programme:** as part of Coca-Cola's '5by20 initiative', the ENGINE programme sought to support both in-school and out-of-school girls (OSGs) in four states of Nigeria to acquire new skills and increase their incomes as a result. This was designed to be achieved by offering specialised tutorial courses to groups of in-school and out-of-school girls through different components.

- **Discovery Learning Alliance's project:** working in schools across Northern Ghana, Northern Nigeria and Kenya, Discovery Communications sought to improve education for girls by training teachers to improve the teaching methods they use, and providing educational videos (and necessary hardware) for them to use as part of their curriculum. Additionally, Discovery sought to change attitudes about girls' education through nationwide talk shows highlighting the value of education for girls.

- **Avanti's iMlango project** aimed to improve marginalised girls' ability to go to school and learn across four counties in Kenya. The iMlango project sought to do this by introducing information and communications technology (ICT)–enabled learning tools to strengthen teaching, and by providing financial stipends to incentivise attendance, develop life skills and build the capacity of teachers to use ICT.

- **Ericsson's Connect to Learn project in Myanmar:** the project aimed to increasing girls' retention in school and improve girls' ability to learn school subjects taught in English including literacy and numeracy. It sought to do this through the provision of broadband, teacher computers, student tablets, a life skills programme (provided by UNESCO), an English language programme and a scholarship programme.

!

*This report includes endline data and findings from three SPW projects: Coca-Cola, Avanti Communications and Discovery Communications. **We did not include the fourth SPW project, Ericsson's Connect to Learn project in Myanmar,** as it had not finalised its endline report and data analysis at the time of writing this report[30].*

Table 1 below is a summary of the SPW projects, and a map of the projects' targeted locations for Discovery, Coca-Cola and Avanti.

Table 1: Summary of SPW projects

	Avanti	Discovery	Coca-Cola	Ericsson *(not included in the EM endline evaluation report)*
Countries	Kenya	Nigeria, Ghana, Kenya	Nigeria	Myanmar
Girls Age /Grades	Primary (all school)	Primary (all school)	Secondary (16-19)	Secondary (all school)
Number of intervention beneficiaries[a]	56,561 in-school girls	302,350 in-school girls	21,162 girls (14,034 out-of-school girls ,7,128 in-school girls)	11,432 in-school girls
Total Budget[b] (of which) DFID's matched funding*	£27,528,066 £14,141,627	£23,957,922 £12,287,098	£6,884,901 £3,164,120	£7,925,227 £3,917,830
Project timeframe	May 2014 – Mar-2017	Sep-2013 – March 2017	Oct-2013 – March 2017	April 2014 – March 2017
Consortium members	sQuid, Whizz Education, Camara Education	Discovery Learning, Alliance	Mercy Corps, d.light Solar Designs	UNESCO, Earth Institute, EduEval, Qualcomm, Finja Five

Data Sources: [a] FM data; [b] Project completion reports; with the exception of Ericsson's budget which was extracted from the contract amendment.

Figure 1 provides a view of the projects' geographic areas of intervention. Each project intervenes in a different geographic location except for Kano (Nigeria) where both Discovery and Coca-Cola are implementing their interventions. Figure 2 focuses on Kano district; highlighting two major issues: most of the targeted locations in Kano are within a 10 miles radius, and in two localities – Dala and Ungogo, and both Discovery and Coca-Cola projects are intervening.

---

[30] The EM was in continuous consultation with the FM to assess whether the inclusion of Ericsson's project was feasible in the endline evaluation. Unfortunately, due to the delay in Ericsson's data analysis and production of the final endline report by its external evaluator, the EM could not incorporate this project in this report. The EM provided continuous updates to DFID about the issue during its bi-weekly management meetings. Prior to making the final decision about excluding Ericsson's project in the EM endline report, the EM provided options to DFID to address this issue. The agreed approach to this issue is set out in Section 2. Due to similar problems at baseline, Ericsson was also left out of the baseline evaluation report.

Figure 1: Locations of projects' targeted areas
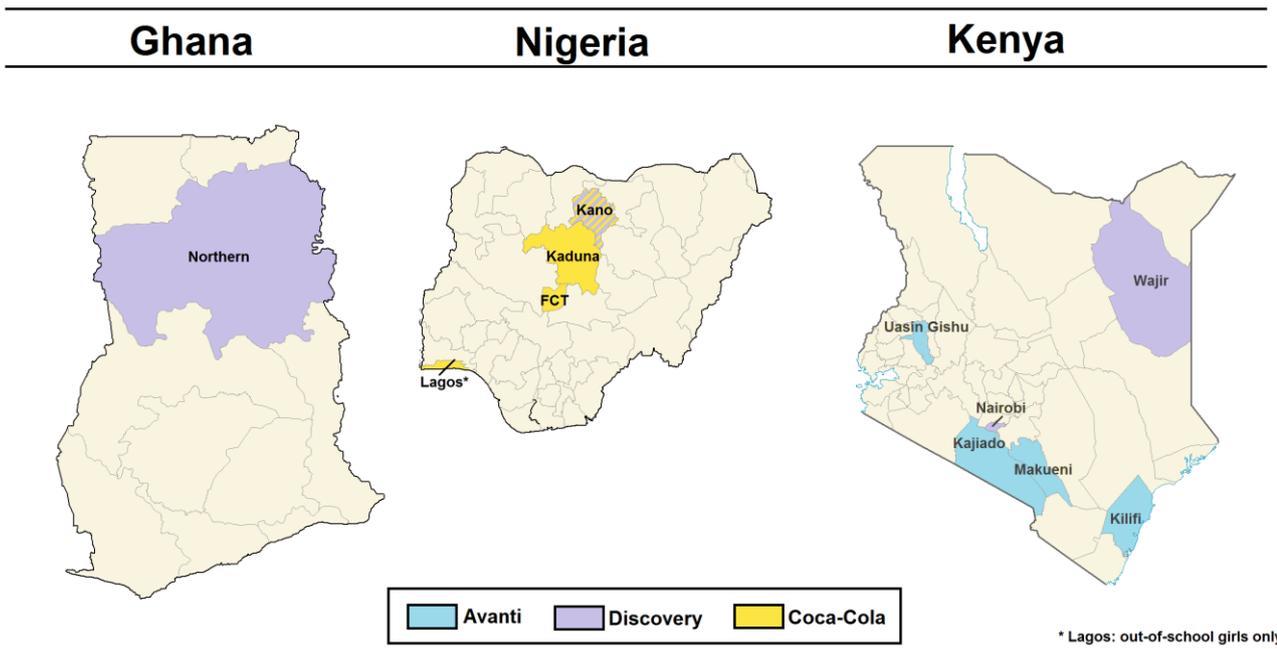
Although both projects target girls of different ages and school levels, questions arise regarding potential overlap of the projects. However, neither project documented any issues, and we do not have any evidence of 'contamination' due to overlap or interference.

Figure 2: Locations of Discovery and Coca-Cola's targeted area in Kano (Nigeria)
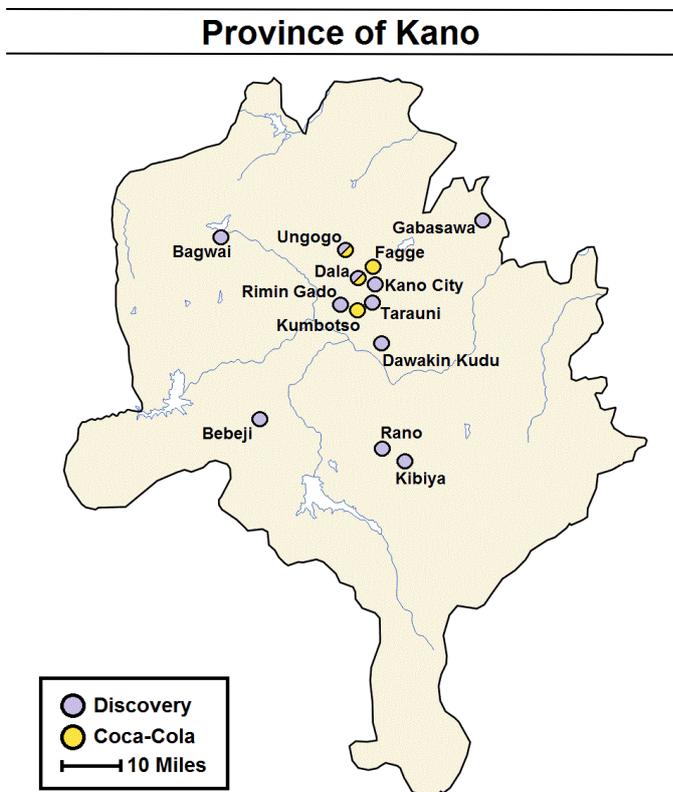
1.1.4   Summary of SPW projects' interventions

Projects adapted specific interventions originating from their own organisational and commercial areas of expertise, and provided those interventions as solutions to tackle some barriers to education, and more broadly, improve the life chances of girls.

To achieve their education-related and broader objectives, projects carried out a wide range of interventions. Table 2 shows that all **projects carried out community-based interventions** to ensure buy-in to their projects, but also to raise awareness about and encourage engagement in education, and girls' education in particular. They also all **carried out extra-curricular activities and non-formal education activities as well as empowerment and self-esteem related activities.** These were mainly in the form of girls' clubs and safe spaces where girls (and in Avanti's case, sometimes boys too) could do academic and non-academic vocational work, receive training on life skills and discuss health, social and other issues (sometimes described as 'sensitive' issues) such as marriage and menstruation. All projects worked with **School Management Committees,** such as Parent Teacher Associations (PTAs), to ensure buy-in to the project, to facilitate the implementation of activities and work collectively to improve the general provision and access to education in the community.

Only Avanti carried out economic interventions to offset the cost of schooling. None of the projects focused on violence as a core intervention in their theories of change.

Table 2: Overview of SPW project' main areas of interventions

Endline evidence by intervention	SPW projects by country and region						
	Avanti	Discovery			Coca-Cola Cycle 1	Coca-Cola Cycle 2	Ericsson[a]
	Kenya	Kenya	Ghana	Nigeria	Nigeria	Nigeria	Myanmar
ECONOMIC — **Economic interventions offsetting the cost of education:** Bursaries, Cash Transfers, Income-generating activities, In-kind support (school kits, menstrual supplies), Loans and savings.	✓						✦
INFRASTRUCTURE & RESOURCES — **Infrastructure and resources for schooling:** School and classroom building/ improvement; Technology in classroom; Textbooks & Learning materials; Toilets & WASH facilities.	✓	✓	✓	✓			✓
TEACHER TRAINING — Teacher training and support: Formal pre-service teacher training; Gender responsive pedagogy; Inclusive classroom strategies; Literacy and numeracy; Peer support and mentoring; Skills training.	✓	✓	✓	✓	✦	✦	✓
COMMUNITY BASED — **Community-based interventions:** Adult literacy; Community gatherings; Household-level visits and support; Media (radio, TV, advertising); Parents' and women's groups; Working with faith groups and traditional leaders; Working with men and boys.	✓	✓	✓	✓	✓	✓	
EXTRA-CURRICULAR & NON-FORMAL EDUCATION — **Extra-curricular activity & non-formal education:** Life skills (incl. SRH); mentoring (e.g. peer support); mixed sex or boys' clubs; non-formal / alternative education; tutoring (e.g. homework clubs); vocational training & economic empowerment.	✦	✓	✓	✓	✓	✓	✦
SCHOOL MANAGEMENT & GOVERNANCE — **School management and governance interventions:** Community and private schooling provision; Technology for school management; Working with local or national education authorities; Working with SMCs, PTAs and other stakeholders.	✓	✓	✓	✓	✓	✓	
EMPOWERMENT & SELF-ESTEEM — **Empowerment and self-esteem:** Activities that promote girls' voice and participation; Mentoring; Role models (older girls, female teachers, parents); Safe spaces.	✓	✓	✓	✓	✓	✓	✦
MARGINALISATION-RELATED — **Marginalisation-related interventions:** Interventions in remote or nomadic locations; Interventions addressing cultural/ linguistic exclusion; Interventions addressing disability Interventions with other marginalised groups.	✓				✦	✦	
VIOLENCE-RELATED — **Violence-related interventions:** Addressing abuse from adults in charge; Addressing child marriage and FGM; Addressing corporal punishment; Addressing peer violence Child protection policies development in schools.		✦	✦	✦			

Source: FM mapping of interventions; Projects' endline reports; Interviews with project staff

Key: ✓ indicates that an intervention of this type is at the core of the projects' intervention strategy. ✦ Indicates that an intervention of this type is used by the project, but is not a core activity.
[a] Ericsson is not included in the EM endline evaluation report.

From the start and throughout the projects' duration, Discovery and Coca-Cola's programme components supporting out-of-school girls were not designed to directly improve their numeracy and literacy. Rather, Coca-Cola focused on improving girls' livelihoods and sources of income, and Discovery focused on improving the life chances of the girls they supported.

All SPW projects combine a range of intervention types. Differences in intervention approaches only emerge when looking at the specific activities that projects are implementing (see Table 3 below).

Table 3: SPW projects' main activities

	Avanti	Discovery	Coca-Cola
ICT equipment and learning aids	**205 schools, provided with:** • Computer labs for individualised learning: 25 computers for each school (average 1000 students per school) • Laptops and projectors • SQUID cards: for monitoring attendance iMlango platform and Maths-Whizz content: Math, English and Life Skills	**1,470 project schools established a Learning Centre (LC) in each school, each provided with:** • 1 large-screen TV • 1 smaller teacher viewing unit • 1 DVD player • a library of over 300 educational video segments	N/A
Teacher training	**3,058 teachers trained in:** • Use of ICT technologies and resources in education	**15,383 teachers trained in:** • Student-centered and activity-based learning • Gender sensitive/responsive teaching practices • Integration of LC materials into daily curricula	**1,472 learning space coordinators trained:** • OSG Learning Space Coordinators (LSCs): 5 days training for OSG LSCs • ISG LSC: 4 days training for ISG LSCs • Training of mentors (1 per 5 Learning Centres)
Community awareness & Engagement		• Community sensitisation workshops • Community Action Plans- manage equipment and support Girls education • National Chat show: 26 episodes in English nationally; and in Hausa for those living in Northern Nigeria	• Sensitisation and mobilisation of community members (SBMCs) and parents • Working nationally with the Central Bank of Nigeria (CBN) to roll out financial literacy curriculum
Girls Clubs/ Safe spaces	• 387 child clubs (209 of which are girl-only)	• 927 girls' clubs have been formed	• ISG training in the safe spaces (38% of total girls) • OSG training in the safe spaces (62% of girls)
Financial Support	**Stipends** Selected based on attendance (less than 70% attendance) validated by School Committees. 10% selected by schools with 10% boys.	N/A	• 1,747 ISGs involved in saving groups activities • 7,473 OSGs involved in saving groups activities • 6276 girls enter value chains (Coca-Cola; d.Light) girls over 18 years old

Discovery and Avanti focused on the use of ICT equipment and digital content as learning and teaching aids. They also provided teacher training. Discovery particularly focused on the pedagogy of teaching, teaching best practice, and gender-responsive teaching. Avanti mainly focused on using equipment and resources as teaching and learning aids for teachers as well as individual students.

Coca-Cola provided different interventions for their in-school and out-of-school target girls. The in-school component of the project enrolled girls in 'Safe Spaces' linked to their school where a programme of tutoring, life skills education and financial education was delivered. The out-of-school programme reached a larger number of girls. This involved enrolling girls into 'Safe Spaces' programmes and delivering content around business and value chain development. Because these interventions were rolled out through the 'Safe Spaces' that Coca-Cola set up, teachers and instructors trained to operate in these spaces were called Learning Space Coordinators (LSCs). In some cases, LSCs were school teachers, and the training they received could have improved their teaching in school, but there is no data available on the background of the LSCs.

Avanti provided stipends mainly for girls, but also for boys based on their (poor) attendance rates.

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 B 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 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 B shows the M&E activities carried out by the FM in the GEC.

Each of the SPW projects was 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, but projects could choose whether or not to follow the recommendations made by the EM about their household surveys. Unfortunately, the flexibility given to projects in designing their surveys, choosing their learning tests and their sampling designs led to high levels of inconsistency between projects. The lack of consistency across the portfolio made it difficult for us, as the EM, to aggregate project datasets, conduct meta-analysis and make comparisons about the impacts and effectiveness of the SPW.

1.2.2 Purpose of the GEC endline evaluation

The overarching purpose of the GEC Evaluation Strategy is to produce reliable evidence of the programme's effectiveness and impacts that DFID, the FM and projects can use to inform improvements during the programme's lifetime, as well as future programme design. In particular, and for the purpose of this endline report, 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, GEC Transitions (GEC-T), and its accompanying 'Leave No Girl Behind' Window that DFID announced in July 2016.

DFID envisaged that the programme evaluation should generate transferable lessons in delivering girls' education outcomes for a wider audience including its partners, governments of GEC countries, and other policy-makers. These lessons would provide insight into what works, what does not, and where and why issues arose in the projects. 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 dissemination plan for the findings of this endline evaluation that will target key education partners of DFID, such as UNICEF, UNGEI, USAID, GPE, UNESCO and the World Bank. Activities in the dissemination plan will be delivered in 2017 following the publication of this report and before the end of the EM's current contract in January 2018.

This endline evaluation also serves an important accountability purpose by providing reliable information about the effectiveness and impact of the SPW projects during the 3 year implementation period.

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

1.2.3   Communication Plan

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

- Consultation and feedback from the FM as part of the validation process including written comments and a group meeting involving DFID (August 2017);

- Sending the final SPW Endline Evaluation Report to SPW partners prior to publication (October 2017);

- Publication of the SPW Endline Evaluation Report by DFID at https://www.gov.uk/guidance/girls-education-challenge (October 2017);

- The EM will produce an Evaluation Briefing Note (October 2017) that summarises the key findings from the report in two to four pages for publication by DFID at https://www.gov.uk/guidance/girls-education-challenge (November 2017);

- DFID will use the Evaluation Briefing Note to raise awareness of the evaluation report and results among DFID's priority audiences for this evidence (November 2017 – January 2018 and beyond);

- We will liaise with the FM to arrange a webinar presentation of the key evaluation findings to the SPW partners (November 2017);

- Face-to-face group presentation of the key evaluation findings to the wider FM team (November 2017); and

- Face-to-face group presentation /seminar of the key endline evaluation findings from the SPW, IW and SCW that is open to all DFID departments and teams (January 2018).

A similar set of communication activities were delivered to disseminate the results from the SCW and IW Midline Evaluation, which proved effective. Prior to completion of this report, together with the FM, DFID and three GEC projects, we presented summary findings that have emerged from this endline evaluation and the endline evaluations of the SCW and IW at the biannual UKFIET conference[1] in September 2017.

1.2.4   Scope of the GEC SPW endline evaluation

The endline evaluation aims to answer the following questions:

- What are SPW projects' target populations? To what extent have target girls and their communities been reached by their interventions? (Section 3.1)

- To what extent has the SPW improved girls' learning and attendance? (Sections 3.2 and 3.3)

- To what extent were SPW interventions addressing key barriers to girls' education and with what effect? What type of intervention works, in what context, and for whom? (Section 3.4)

- How scalable and sustainable are the activities funded by the SPW? (Section 3.5)

The GEC endline 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 the delivery of GEC activities is covered through the FM's performance reporting, and is not within the scope of this evaluation. A discussion of progress against output targets can be found in each project's Endline Evaluation Report.

The GEC endline evaluation focuses on answering questions about the effectiveness and impact of the SPW projects as well as the potential sustainability of the activities delivered – our assessments of sustainability rely on data and information provided in project evaluation reports. As far as possible, the findings, conclusions and recommendations produced in this report reflect the unique characteristics of these private sector partnerships, but specific findings and lessons learned about the process of setting up and delivering these partnerships in the SPW are covered in the EM's Process Review[31] and in the FM's Strategic Partnerships Consultation[32]. The FM's Strategic Partnership Consultation Report provides valuable insights into the original ambition for the SPW and

---

[1] UKFIET, The Education and Development Forum – https://www.ukfiet.org/conference/
[31] Coffey (2015), GEC Process Review Report
[32] PwC (2015) Strategic Partnerships Consultation

what had worked well and less well with regards to the process of developing the partnerships at that stage. For the purpose of putting this evaluation in context the report usefully describes DFID's ambition as follows:

> *"The Strategic Partnerships Window was intended to work with new partners, in new ways, in multiple countries, with exacting M&E standards and with limited time for design. The level of ambition has meant that challenges were inevitable, but good progress has been made. The balance of the 'partnership' versus the 'project' benefits continues to shape decision-making, engagement and support to the Lead Partners."*

It is important to recognise the context in which these partnerships were built. These types of public-private partnerships were in themselves new and untested. Private sector partners were expected to come with solutions that could be both commercially sustainable and deliver challenging education outcomes for marginalised girls. The purpose of this endline evaluation of the SPW is to produce reliable evidence of the programme's effectiveness and its impact on the marginalised girls targeted by the three projects. As such, it does not cover the public-private partnership building process that inevitably shapes the type of project designs that were implemented. However, the context in which these projects were designed and delivered should be taken into account when reviewing the findings, conclusions and recommendations arising from this evaluation.

1.2.5   Structure of the SPW endline report

The report is organised around the endline evaluation questions.

Section 2 presents the approach and methodological challenges to synthesising and aggregating findings from the endline research conducted by the SPW projects and further analysis undertaken by the EM using the projects' datasets.

Section 3 focuses on key findings, i.e. the extent to which target girls and their communities have been reached by SPW interventions, the extent to which SPW projects improved girls' attendance and learning, and which type of interventions work best, in what context, and how.

Section 4 and Section 5 present our conclusions and recommendations for the different audiences of this evaluation.

A list of annexes can be found at the end of this report, and comprises:

- **Annex A.1:** GEC Evaluation Manager Terms of Reference;
- **Annex A.2:** GEC Theory of Change;
- **Annex B:** Roles and responsibilities for the GEC Evaluation;
- **Annex C:** SPW projects' ToCs and Intervention Mapping;
- **Annex D:** Learning Outcomes;
- **Annex E.1:** Effectiveness Tables
- **Annex E.2**: Sustainability Tables; and
- **Annex F:** List of references.

2 Evaluation Approach and Methodology

2.1 Overview of the GEC Strategic Partnership Window evaluation strategy

For the Strategic Partnerships Window (SPW), the GEC evaluation strategy focuses on analysing evidence and data solely produced by the projects and by the project-led evaluation activities.

Projects assess the impact of their interventions on their target groups (see Section 3.1.1. for an overview of each project's target groups). They 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 adaptation and improvements to their project designs. The main data sources comprise (detailed below): Projects' Endline Evaluation Reports; Projects' Outcome Spreadsheets; Projects' Datasets; Projects' periodic reports; and interviews with project staff.

The Evaluation Manager (EM) conducted a meta-analysis of project-level evaluation findings to assess the overall impact of the GEC SPW interventions on girls targeted by SPW projects. In this report we are not evaluating the inherent benefits arising from the partnership between the public and the private sectors. We set out to produce lessons learned to inform GEC and wider DFID programming, and to build the wider knowledge base around what works in girls' education through the type of public-private partnerships fostered in the SPW.

2.1.1 Project evaluation design and changes since baseline

Support provided to SPW projects at endline

The GEC Fund Manager (FM) supported projects with their evaluation and monitoring systems on a continuous basis (refer to Annex B for roles and responsibilities). The FM advised projects on changes to their evaluation design, and the adequacy of learning tests and analytical models. The FM also led the quality assurance of projects' research instruments and evaluation reports.

The EM provided the following support to projects at endline:

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

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

- Provided guidelines for endline analysis and reporting;

- Reviewed projects' data collection tools for endline (both qualitative and quantitative);

- Provided feedback jointly with the FM during calls with the projects and their external evaluators; and

- Reviewed and provided feedback on projects' endline evaluation reports jointly with the FM.

Box 1: Projects' impact evaluation design

Initially, for each of the SPW projects, external evaluators planned to conduct randomised control trials (RCTs) as an appropriate method for addressing their evaluation questions. However, due to their sampling approaches and to follow the same protocols for data collection in treatment and control groups, they actually conducted quasi experimental (Q-E) evaluation designs. Indeed, school assignment was not totally random because the projects had to purposively select (to some extent) the school/areas they intervened in, and then identify control groups that were comparable to their treatment schools/areas. It is worth noting that in all the SPW project reports, the evaluation design used is still called a 'randomised control trial', instead of 'quasi-experiment' (as we designate it in the present report).

Changes since baseline and data quality at endline

SPW projects assess the impact of their interventions on their specific target groups. Projects generate findings about what works, what does not (and why) at the project level, draw lessons learned about their theories of change, and reflect on possible improvements to their project design. The project-led evaluations included the following activities:

- Commissioning an independent evaluator to collect data at baseline and endline, and produce an evaluation report at each stage that complies with a template provided by the FM and EM. For Discovery and Coca-Cola projects, external evaluators also carried out 'light touch' midline evaluations based on qualitative data. The purpose of these evaluations was to generate learning from projects, and adapt their interventions based on that. The EM and FM were not significantly involved in providing support for the projects' midline evaluations;

- At baseline and endline, collecting a combination of quantitative and qualitative data in intervention and control communities (or schools), including a longitudinal household survey. SPW projects were encouraged to use the standardised survey template provided by the EM to collect data consistently across the window, but had the flexibility to design their own survey templates;

- At baseline and endline, testing literacy and numeracy using standardised international tests, and conducting appropriate statistical analysis to report on changes in learning outcomes; and

- Reporting of outcome evaluation findings using the GEC reporting templates (report and outcome spreadsheet).

Following the GEC baseline research, all SPW projects have sampled control areas except for Discovery in Wajir (Kenya). Comparative impact evaluation of the project was undertaken in all other locations, using a quasi-experiment (see Box 1 above and Table 4 below) except for Coca-Cola's Cycle 1 which did not carry out any statistical analysis between treatment and control groups at baseline and endline.

Table 4: SPW projects' evaluation designs and data quality at endline

Project evaluation designs and data quality at endline	SPW projects by country and region								
	Avanti			Discovery				Coca-Cola	
	Group A	Group B	Group C	Wajir [c]	Nairobi	Ghana	Nigeria	Cycle 1	Cycle 2
	Kenya			Kenya				Nigeria	
Evaluation designs [a]	Q-E	Q-E	Q-E	PE	Q-E	Q-E	Q-E	N/A	Q-E
Control group data	✓	✓	✓	Ø	✓	✓	✓	✓	✓
Data quality for literacy and numeracy scores [b]	2	2	2		3	3	3	1	1
Data quality for attendance rates [b]	1	1	1		3	3	3	2	2

[a] PE: Performance Evaluation; Q-E: Quasi-Experimental design.

[b] Data quality defined on a scale from 1 to 3: 1 – Inconclusive/major issues; 2 – Partially conclusive; 3 – Conclusive. This relates to whether the tests were correctly administered and the data clearly entered /labelled as submitted to the FM and the EM. These scales are adapted from the FM appreciations and comments provided to the EM on the projects' data quality. As such, the data might be of adequate quality and still prevent any conclusion on impact if other issues arise (sampling, statistical analysis, etc.).

[c] For Wajir (Kenya), where a performance evaluation has been conducted, the FM did not provide any evaluation of the project's data quality neither for learning, nor for attendance.

	Change in evaluation design since baseline

Changes in evaluation design and data collection

The EM faced the following challenges when carrying out its analysis and synthesis of the projects' data due to the changes in evaluation designs and data collection by projects:

- **Discovery compared two streams of implementation separated by one year in Wajir whereas the RCT design that it used in other places measured two years of activity.**

  In Kenya, where Discovery was planning to deploy the programme in Nairobi and Wajir areas, the study was split into two parts during the Inception Phase. In Nairobi, treatment and control groups were selected and data collected at baseline and endline to perform impact analysis. During the design phase of the

impact evaluation, stakeholders agreed not to include project schools in Wajir (Kenya) due to the very limited number of project eligible schools for treatment and control groups in these areas. As a compromise, Discovery designed a performance evaluation using a pipeline approach (also called a stepped-wedge design). The design exploits the variation in the timing of programme implementation and uses beneficiaries chosen to participate in the project at a later stage as the comparison group. It assumes that the two groups are similar and therefore comparable in terms of the (education) outcomes. The programme of activities was first rolled out across one group of schools (i.e. Year 1 schools). After a year of implementation, learning and attendance was tested for Year 1 schools, and for another group of different schools – (Year 2 schools) that had not yet received the intervention. Another round of data (Round 2) was collected a year after the first round was collected after Year 1 schools had been benefitting from the intervention for two years, and Year 2 schools for one year. As a result of this approach, there is only one year between Round 1 and Round 2 data collection points whereas endline data has been collected two years after baseline for the other areas evaluated through a quasi-experimental design. This makes it difficult to reasonably compare the impact and effectiveness of activities delivered in Wajir compared to other places where Discovery worked.

- **In Cycle 1, Coca-Cola's new external evaluator (after baseline) was not able to match the cohort at endline with the cohort used at baseline, which meant longitudinal statistical impact analysis could not be undertaken. Moreover, for Cycle 2, the FM assessed Coca-Cola's literacy data as unusable and it has not been analysed in this report.**

  Coca-Cola changed its external evaluator after Cycle 1 baseline research and instead commissioned the same evaluator that was appointed for Cycle 2 baseline and endline evaluations. This new external evaluator experienced difficulties in matching observations for Cycle 1 from baseline and endline data for learning and did not conduct any statistical comparison between treatment and control groups at baseline and endline that would permit measurement of the project's impact. From the FM's perspective, baseline and endline scores for literacy and numeracy are still comparable. Accordingly, we proceeded to an analysis of Cycle 1 average scores for literacy and numeracy. However, as baseline and endline samples could not be matched by the external evaluator, we were not able to draw conclusions about the statistical significance of Coca-Cola's impact on learning. In Cycle 2, data quality for literacy was assessed as unusable by the FM, and as a result is not presented in this evaluation report. We have only analysed the project's impact on numeracy in this cycle.

- **At endline, for both Avanti and Coca-Cola, attendance data measured for treatment groups could not be compared to control groups**.

  Avanti's electronic attendance monitoring system provided very accurate data about the percentage of school days attended by the cohort girls, but the poorly kept attendance records in control schools did not provide comparable evidence. Therefore, we could analyse the change in attendance from baseline to endline in the treatment group, but we could not conduct a counterfactual analysis comparing this to the change in the control group. In both Cycles 1 and 2, Coca-Cola measured the average monthly attendance rate in the Safe Spaces supervised by the project and not in schools – this did not allow any comparison with a control group but allows a comparison between in-school girls and out-of-school girls and between cycles.

- Finally, we also noted that while Avanti tested both Grade 1 and Grade 3 students at baseline, the **Grade 1 cohort was dropped from the evaluation cohort at endline** because the probability of ceiling effects was high. Therefore, only girls that were in Grade 3 at baseline were tested at endline.

It is important to note that there is a risk of penalising a project that is able to produce better quality data and endline evaluation reports, such as the Discovery project which provided an extensive amount of evidence and good quality data. We were able to carry out more analysis with Discovery's data than we were able to for other projects, and Discovery's endline evaluation report provided more granular evidence about what did not work compared to the other SPW projects, which does not necessarily mean the project did not work as well, but is more likely a result of a reporting deficit in the other evaluation reports.

2.1.2 The EM evaluation approach

As the EM for the SPW, we assessed the overall impact of projects on targeted girls in treatment communities[33]. Our evaluation design relied solely on SPW projects' reported evidence, analysis and data.

We therefore reviewed and carried out a **meta-analysis of SPW reports, datasets and outcome spreadsheets** submitted by SPW projects. The process adopted aimed to synthesise the evidence provided by projects to enable us to report on evaluation findings at the SPW-level across all three partners.

Our approach also involved assessing the quality of the data and findings against the quality assurance criteria defined by individual projects, the FM and the EM (Figure 3).

The step-by-step process adopted is as follows:

- Mapping of project documents and data available (Project Endline Evaluation Reports, Outcome Spreadsheets, Project Datasets; Project Baseline Reports, Project Proposals; Project M&E Frameworks; and projects' periodic reports including annual reviews and completion reports);
- Systematically extracting the data and analysis from project documents, including Project Datasets;
- Ensuring the consistency and quality of reported findings;
- Synthesising the evidence base reported by projects at endline;
- Carrying out semi-structured interviews with project staff to unpack detailed activities to enable us to interpret reported findings (four interviews with each of the projects); and
- Conducting analysis that responds to the endline evaluation questions.

Figure 3: Meta-analysis and synthesis approach

Project Endline Evaluation Reports
Outcome Spreadsheets
Project Datasets
Interviews with staff

Data extraction and reanalysis

Data quality assessment

Meta-analysis/ Synthesis/ Triangulation

Evidence used to answer endline evaluation questions

Approach to synthesising SPW projects' findings

For the evaluation of the SPW, the EM did not collect any additional data beyond the data collected by projects and their evaluators. Therefore, the findings in the report are not intended to supersede the findings of individual projects but rather to synthesise evidence across the funding window. The goal of this synthesis is to apply a common evaluation framework to all projects, measuring impact on the same outcomes and allowing a comparison across the window of project approaches.

> *It has to be noted that projects did not conduct (or report on conducting) any process evaluation in relation to their implementation processes. This rather limits the ability to assess whether project effectiveness (or lack of effectiveness) relates to design and/or implementation strengths and weaknesses.*
>
> *We recognise that **there are many factors (related to both project design and implementation) that could potentially explain why projects have been effective or not**. Where relevant, we have indicated the limitations of our interpretation of projects' findings.*

---

[33] Refer to project targeting (Section 3.1.1).

Our quality assurance role as the EM therefore focused on: (1) considering the comparability of measurement tools used across projects and consistency in reported measures across Project Endline Evaluation Reports, Outcome Spreadsheets and Project Datasets; and (2) assessing the quality of the data collected and reported. Table 5 below provides the list of consistency and quality criteria used to synthesise the SPW project data and analysis.

Table 5: Criteria used for the synthesis

Consistency criteria	Quality criteria
✓ Comparable measurement tools (e.g. learning assessments) ✓ Comparable indexes compiled by projects for reporting on educational outcomes	✓ Defined by each individual projects' external evaluator ✓ Quality Assurance conducted by the EM and the FM prior to Project Endline Report approval

Despite the triangulation of findings across the different sources available and the EM reanalysis of projects' data, there are limitations to the EM's interpretation of the synthesised data.

Changes to the evaluation approach at endline

The EM's evaluation approach has remained largely unchanged since baseline. We used the same types of documents and data provided by projects, but also conducted interviews with project staff (four interviews for each of the three projects included in this report). The process of extracting data from project documents, however, was adapted at endline to incorporate lessons learned from baseline. In particular, a 'data extraction' template was used across projects to ensure consistency, focusing on:

- Changes – What has changed since baseline in terms of barriers to girls' education? What has not changed?
- Linkages to interventions – As a result of which project intervention? How? Where? For whom?
- Effects – What was the effect on education outcomes (attendance, learning)?
- Other contributing factors – What other non-GEC activities have happened and with what effect on barriers? Is there evidence of coping strategies and/ or enabling contexts?
- Micro-level changes – Are there any stories of interest, individual pathways that can shed light on how change happens for girls and their communities for a particular project?

The focus in the process of data extraction was on change from baseline. To some extent, this makes comparisons across projects more feasible than it was at baseline. Because of differences in context and measurement approaches across projects, it was difficult at baseline to make comparisons between projects on many of the educational outcomes and educational barriers. Changes in these factors are more directly comparable, so our analysis at endline could make more direct comparisons across projects.

2.2 Methodology and data sources

The SPW Evaluation Strategy mainly used project-led evaluation evidence and data to assess projects' effectiveness and impacts. This is the same approach that has been used for the GEC's Innovation Window. Additional interviews with the projects' staff were intended to help fill gaps in the information provided in the projects' reports, contextualise and interpret emerging findings and provide an independent source of evidence to compare against qualitative evidence in evaluation reports.

2.2.1 Project information

The GEC Evaluation Strategy required all SPW projects to carry out independent qualitative and quantitative research at baseline (in 2014/ 2015), and at endline (in 2016/2017). Unlike the Step Change Window and Innovation Window, no midline evaluation was required for SPW projects.

All SPW projects' external evaluators conducted surveys using questionnaires and sampling frameworks that were reviewed by the EM and the FM during the development of their M&E Frameworks prior to starting their baseline

research, and again as part of their preparation for endline evaluations. All SPW projects' external evaluators tested the literacy and numeracy skills of in-school girls in target communities, conducted qualitative research and were also encouraged to draw on existing sources of secondary data.

> **!** ***SPW projects could develop or adapt their own qualitative and quantitative research and sampling designs (by using a randomised household survey or by sampling girls at school), and they may have taken different approaches with regards to qualitative sampling or the development of interview guides****. While quantitative data (*Project Datasets*) were shared with the EM along with *Projects Endline Evaluation Reports*, projects were not required to submit their qualitative data to the FM. As a result, the qualitative findings presented in this report are based solely on SPW projects' analysis, as presented in their reports.*

Data sources

The main evidence gathered by projects through their endline research was documented in three different formats, as detailed below. The three sources of information have different strengths and weaknesses.

- **Project Endline Evaluation Reports** present evidence, key findings, and lessons learned based on the data analysis led by projects and by their independent evaluator/ affiliated researchers who had all committed in their M&E Frameworks to achieve: samples that were highly representative of their target populations; statistical power; and analytical quality. The Project Endline Evaluation Reports focused on testing a project's theory of change and assumptions about target groups, educational outcomes and barriers to education. However, reporting against indicators was not consistent across projects and project reports did not always reflect on the range of indicators of interest for GEC endline analysis at the programme level, and as such was **not always in a standard format and respondents did not always answer the same questions** (see Box 2).

- **Outcome Spreadsheets** were used by projects to report the levels of attendance and learning (literacy and numeracy) at endline, which were the primary outcomes for the GEC. They are a way of consistently capturing key outcome data and reporting on progress against targets for learning and attendance for all projects. The Outcome Spreadsheets have the advantage of providing a relatively standard format and allowing disaggregation by grade, subject to some variation in the learning assessment tools used. It is worth noting that none of the SPW projects were subject to Payment by Results (PbR), unlike the other SCW and IW projects for whom part of their payments were based on successfully achieving their midline and endline targets for literacy and numeracy. The Outcome Spreadsheet was critical to the FM's assessment for PbR purposes.

- **Project Datasets** were submitted by projects along with their Endline Evaluation Reports**.** They include the raw data from the household surveys and/or in-school surveys. The EM has carried out an independent analysis of this data for a selected number of key variables where the relevant information was available, documented and comparable. This "reanalysis" (presented in boxes in the Targeting and Reach Section 3.1 and in the Effectiveness Section 3.4) aimed to cross-check and verify the figures and findings presented by the projects in their endline evaluation reports. The EM chose to **focus the reanalysis on changes in barriers and the level of exposure target groups experienced to different types of interventions**, to avoid a duplication with the reanalysis conducted by the FM on learning variables. The process followed by the FM to reanalyse learning outcomes was shared with the EM for verification purposes for each project, along with a summary of key issues found in the datasets during the reanalysis.

Box 2: Limitations to the reanalysis of Project Surveys Datasets and meta-analysis

Project data received by the EM varied in terms of the types of surveys administered, number of surveys administered, survey questions asked, type of respondents, data quality, and approaches to merging the data. This meant that data quality and availability was a major consideration for what kinds of analyses could be conducted to support our evaluation of the projects' effectiveness. We assessed the quality of evidence and conducted in-depth analyses where data would permit and where it could contribute to the most important questions addressed by the evaluation report. The downside to this approach was that we were not able to systematically conduct the same kind of analysis of projects' impacts on all barriers across all projects. As a result, the findings may not always provide a good comparative view. For example, because the structure of Avanti's data was entirely different from

---

Discovery and Coca-Cola's data we conducted many analyses that compare the Coca-Cola and Discovery projects directly but look at Avanti separately when data was available.

Changes in survey instruments from baseline to endline

In many cases, projects made changes to the questions or design of their survey instruments between baseline and endline. Ideally, we would like to have been able to directly analyse changes in the same indicators from baseline to endline across all projects. The changes made meant that sometimes we had to either compare responses to modified questions or we were not able to track changes over time at all.

Inconsistencies in survey instruments used across projects

While Coca-Cola and Discovery based their household surveys on the template provided by the EM, Avanti used an entirely different instrument, so in most cases, it was not possible to directly compare indicators of different barriers in Avanti's project with those in Coca-Cola's and Discovery's projects. For individual questions, the alterations made by different projects to the template meant that there was a relatively small number of survey questions that we could directly compare across projects.

Poor documentation of attrition[34]

With the exception of Discovery, there was very little documentation available about the reasons for attrition from the sample, especially in the household survey. With large attrition rates, and possibly undocumented substitution households, we do not have a good way to assess how changes in the sample composition might account for apparent changes from baseline to endline. This was especially problematic for Coca-Cola. Our results find many puzzling changes from baseline to endline for Cycle 1 over a period of 9 months, which are likely due to changes in the composition of the sample that have not been properly documented.

Missing data

Within datasets, many variables showed a significant non-response. Without good documentation of why there are missing values for many variables, it is possible that this could introduce unknown biases into the analysis.

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

In this report, we present a synthesis of projects' findings about barriers and assess the extent to which these findings are being supported by projects' endline evidence.

Projects identified specific barriers at baseline that they assumed drive educational marginalisation in the target areas. While some of these barriers are structural or environmental and beyond the projects' direct control (such as the occurrence of droughts or political violence), others may be tackled through targeted interventions and support (such as negative attitudes towards girls' education or a lack of adequate sanitation facilities in schools).

It is important to note that this report presents evidence collected by projects of the **most reported barriers** *perceived* to be preventing girls from attending school and learning. As such, barriers may not be *actual* barriers (e.g. fear that violence may occur on the way to school *versus* reports of violence occurring on the way to school) but the influence of these **barriers, either actual or perceived**, is assumed here to similarly prevent girls from attending school and learning.

In the EM baseline report, barriers were categorised into key thematic areas based on a synthesis of the projects' baseline evaluation reports. The categories were: poverty; school-related factors; girls' aspiration, decision-making and early marriage; negative attitudes; and violence. In this report, we use the same categories to assess the evidence that there were changes in these barriers from baseline to endline and that projects contributed to these changes.

We follow a four-stage approach to assessing the effectiveness of interventions. The first three stages are a process of evaluating and synthesising findings from the evaluations of each project; the last stage is a triangulation with our reanalysis of the projects' data. This synthesis process is described below.

---

[34] We chose not to investigate further the potential effects of attrition on learning outcomes in SPW projects as there are no robust findings of projects' impacts that derive from the quasi-experiments.

1. We review the situation at baseline, including which barriers the projects found, and how the projects intended to address those barriers.

2. Evidence from the projects' evaluation reports were reviewed, assessed, and synthesised to determine whether there had been a change in each barrier. Given the imprecise nature of much of the evidence and the lack of reliable metrics for these barriers, changes in barriers are interpreted only as a '**direction of change**', either positive, neutral or negative.

3. Using the same sources of evidence, we synthesise project evidence on whether any changes in barriers can be attributed to the SPW projects. This draws both on quantitative measures presented in project reports and on qualitative evidence such as beneficiary interviews in the reports. While this process relies heavily on the evidence presented by the projects, we use the same standards of evidence across all projects.

4. Using the raw data provided by projects, we conduct independent analyses to validate findings from steps two and three or address the questions from another perspective. This stage is highly dependent on data availability and quality, so we only move to this step for the most important questions where data is available.

2.2.2 Secondary data

The interpretation of projects' endline findings was supported by the use of secondary data. More specifically, we mainly gathered information from international sources of secondary data relating to girls' literacy and attendance.

Table 6: Secondary data sources used

Educational outcomes
Learning
To analyse the change in reading fluency from baseline to endline across the SPW projects, we used the benchmark proposed by Abadzi[35], which provides the **minimum average words per minute in a given grade**. We also used the evidence highlighted by the author that below 45 words per minute, a student is not able to understand a simple paragraph of text.
Attendance
Primary school female national net attendance ratio for 2013-2014 refers to the number of children attending primary school[36] who are of official primary school age, expressed as a percentage of the total number of children of official primary school age. This data is the most recent data available collected by Demographic and Health Surveys (DHS).This is the most recent data available for attendance on the World Bank data base.

2.2.3 Triangulation and synthesis process

As described in Section 2.1, this report draws evidence from a range of different data sources to answer the SPW Evaluation Questions. Table 7 below presents a simplified version of the GEC Evaluation Framework showing how methods and data sources have been triangulated to answer each evaluation question.

---

[35] 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.

[36] Attendance is defined in a different way in the GEC (percentage of school days attended by enrolled girls). As such, the UNICEF definition is closer to the definition of an enrolment rate.

Table 7: Overview of the streams of evidence used to inform the analysis presented in this report

GEC Evaluation Questions	Project information					Secondary data
	Project Endline Report	Project Dataset	Outcome spread-sheet	Projects report	Interviews	
Relevance: To what extent has the GEC reached marginalised girls? (Section 3.1)	✓	✓		✓	✓	
Impact: What impact has the GEC had on marginalised girls' learning? (Section 3.2)		✓	✓			✓
Impact: What impact has the GEC had on enabling marginalised girls to be in school? (Section 3.3)		✓	✓			✓
Effectiveness: What has worked, why and with what effects? (Section 3.4)	✓	✓		✓	✓	✓
Sustainability: How scalable and sustainable are the activities funded by the GEC? (Section 3.6)	✓					✓

The GEC evaluation was **not designed to test and compare the effectiveness of specific intervention types**, or to assess what type of intervention is most effective, except for Avanti to some extent (see Box 3). Projects carried out a variety of activities in their intervention areas and it was not possible to single out the impact of one specific intervention using quantitative methods.

Box 3: Avanti's specific intervention design

Avanti designed its project evaluation to compare the effectiveness of different combinations of specific intervention types. The cohort of 260 schools tracked by the project were broken down into four groups (A, B, C and D). Groups A, B and C received different combinations of interventions, and Group D was the control group. The sample size for each of these four groups is similar[37].

As shown in Table 8, Group A received the whole intervention, whereas Groups B and C only received some components provided to Group A. Group B did not received personal tutoring in maths and ICT labs, and Group C did not receive the stipends component.

**Table 8: Avanti's intervention for each of the treatment groups**

Avanti's programme components	Group A	Group B	Group C	Group D
	Full Intervention	No Personalised Tutoring	No Stipends	Control Group
Access to iMlango learning platform	✓	✓	✓	
Maths Whizz Personal tutoring	✓		✓	
Maths Whizz whole class teaching	✓	✓	✓	
Payment of stipends and prizes	✓	✓		
Attendance monitoring	✓	✓	✓	
Internet Connectivity	✓	✓	✓	
ICT suite (15-25 computers)	✓		✓	
Computer & Projector (for whole class teaching)	✓	✓	✓	

Some of the components of Avanti's programme, such as access to iMlango leaning platform, Math Whizz whole class teaching, as well as the provision of computers and projectors for class teaching are provided to all three treatment groups. Comparing A to B or C provides an estimate of the additional impact of personalised tutoring or stipends conditional on getting the rest of the interventions, but does not provide evidence on what the impact of these components would be without the rest of the intervention package.

---

[37] This means that, when an overall comparison of groups A, B and C against group D is done, control group is very small compared to the treatment one, which does increase the standard error in difference estimator, and reduces the power of the significance test compared to a test with a control group size similar to the treatment one. However, Avanti's goal was to compare each variant of the treatment to the control group. Nevertheless, the size of these samples is smaller than the other projects', so the significance test has less power for Avanti.

Therefore, using an **iterative approach** to synthesising evidence (Figure 4), we drew upon the analysis presented in the projects' endline evaluation reports, the outcome data on learning, attendance and enrolment, the findings from projects' dataset reanalysis and the secondary data available – to unpack how and why changes have (or have not) come about, and how well different types of interventions have delivered their intended outcomes.

Figure 4: Triangulation and synthesis process
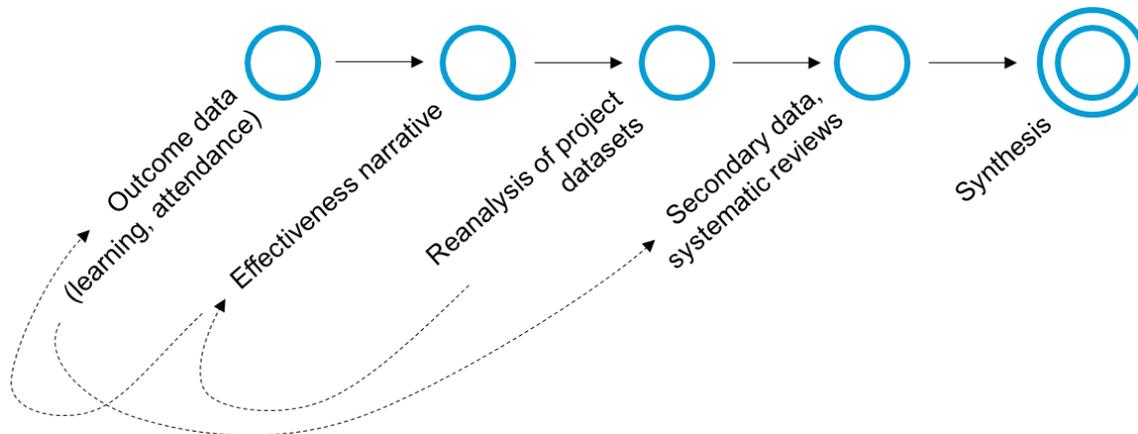
2.3 Methodological limitations and mitigation strategies

2.3.1 Limitations of the EM's endline approach and mitigation strategies

Challenges identified by the EM while extracting, analysing and synthesising the data are listed below:

- **Significant gaps and quality issues with the evidence base:** significant gaps and weaknesses in the evidence available arose in relation to some of the key GEC outcomes. Missing and/or unreported figures, contradictory values reported in the Project Endline Evaluation Reports and other inconsistencies in Project Datasets were addressed where possible by triangulating the available evidence (e.g. Outcome Spreadsheets). To give an example of data quality issues: it was not possible, for the majority of projects, to conduct an analysis of learning outcomes across subgroups (e.g. rural/urban populations, disabled groups, socio-economically disadvantaged groups) or for groups facing specific barriers (e.g. poverty, violence, early marriage). One of the problems we faced when considering this kind of analysis was that most of the demographic information or variables that would identify subgroups was collected as part of a household survey whereas the learning data was collected in a separate survey of girls in school. For Avanti, these surveys did not attempt to follow the same girls. For Discovery and Coca-Cola, the learning surveys and household surveys were not linked.

- **Contradictions arising from a diversity of data sources:** the diversity of data sources and different types of data reported by projects (quantitative versus qualitative data, population of reference, reporting style etc.) led to difficulties in synthesising the findings that emerged into a coherent narrative. However, the structured and systematic approach used for the analysis, triangulation and synthesis of the data helped resolve contradictions arising from the analysis by providing a transparent means of explaining why they occurred.

- **Synthesis challenges:** a carefully structured approach to the synthesis of project findings was adopted in order to mitigate against the effects of different types of bias. Challenges identified include:

  - potential **sources of heterogeneity**, including project research methodologies, the narrative versus quantitative nature of the synthesis, degrees of data validity, cultural sensitivities and contextual factors; and

  - the identification of **adverse synthesis effects** – effects that were identified as very likely to have been lost during the synthesis process; for example, if two equally valid sources of data (such as Project Endline Report findings and Outcome Spreadsheets) presented different findings, there was a tendency to conclude that this was an inconclusive finding, leading the EM to investigate a third source (e.g. Project Dataset).

- **Interviews were limited in interpreting findings:** As mentioned, the EM carried out four interviews with staff from each of the three projects. The interviews with project staff aimed at providing detailed information regarding implementation of interventions as well as provide a chance for projects to provide contextual background and possible explanations for emerging findings. However, interviewees did not always have detailed knowledge of project implementation and sometimes provided more general evidence of project successes that was not easy to verify. Additionally, some interviews were crowded, which hindered the ability to have a more in-depth and meaningful discussion to reflect on lessons learned, particularly with regards to 'what did not go so well'.

We do not expect these limitations to the approach to significantly compromise the quality of the synthesis of the endline findings, or its capacity to add significant value to DFID's understanding of how and to what extent the SPW projects managed to improve girls' educational outcomes.

2.3.2   Limitations of the SPW projects' endline research

At endline, SPW projects faced issues with timing, attrition, inaccurate learning assessment scores, matching of cohort observations, and comparison of data across baseline and endline, the usability of school register data, and contamination issues. This section summarises some of the key challenges and mitigation strategies reported by projects in their endline evaluation reports.

Project data collection timings and length of implementation

The period between SPW projects' Inception Phases during which baseline research was conducted and projects' data collection at endline is not uniform across projects – implying that projects measure impact over longer or shorter periods of implementation.

Table 9 shows that there are discrepancies across projects. Between baseline and endline, the number of school terms (three months) varies from one project to another. The minimum is three terms for Coca-Cola, the nine month education cycles covered three school terms. For Discovery, the duration between baseline and endline was five terms in Nairobi (Kenya) and Ghana, and four terms in Nigeria (cohort girls were in Grade 5 at baseline and in Grade 6 at endline), whereas for Avanti, it was six terms (cohort girls from Grade 4 at baseline were in Grade 6 at endline). We can also note that despite the three terms indicated for Discovery in Wajir (Kenya) at Round 1, Year 1 school girls had already received the intervention during one year (three terms), and the whole project duration is similar to the other areas covered by Discovery. This suggests that the **impacts on girls' educational outcomes may not have been realised at the time of endline data collection** for some projects.

Table 9: Project implementation and data collection timings

Project data collection timings	SPW Projects by country and by region						
	Avanti	Discovery				Coca-Cola[c]	
	Groups A, B and C	Wajir [a]	Nairobi	Ghana	Nigeria	Cycle 1	Cycle 2
	Kenya	Kenya				Nigeria	
Start of project's implementation	May-14	May-14 (Year 1) May-15 (Year 2)	May-14	Apr-14	Oct-14	2014	2015
End of project activities	March-17	Sep-16	Sep-16	Sep-16	Dec-16	2015 (9 months)	2016 (9 months)
Baseline data collection start date	Aug-14	Mar-15	Mar-14	Mar-14	May-14	n/a	n/a
Endline data collection start date	Oct-16	Feb-16	Feb-16	Feb-16	May-16	n/a	n/a
No. of terms between baseline and endline	6	3	5	5	4 [b]	3	3
Grade at Baseline	Grade 3	Grade 5	Grade 4	Grade 4	Grade 4	Grades 7-11	Grades 7-11
Grade at Endline	Grade 6	Grade 6	Grade 6	Grade 6	Grade 6	n/a	n/a

[a] For Wajir (Kenya), we filled the baseline and endline cells with information for Round 1 and Round 2.

[b] The 2014/2015 school year was shortened due to a confluence of events such as a presidential election and Ramadan falling in the summer. As a result, the school year was extended through August although the number of weeks school was in session was still less than planned.

[c] Exact dates are not provided in the documents the EM accessed.

Attrition

Attrition was high across the window (see Table 10). Half of the sample has been lost in Avanti's sample and Coca-Cola's sample for Cycle 2 projects in the treatment and control groups. In Cycle 1, for which we did not have reliable data for attrition, the whole out-of-school sample has not been re-contacted due to problems with the re-contact protocols used at baseline.

Table 10: Attrition rates among IW projects

Project data collection timings	SPW Projects by country and by region								
	Avanti			Discovery				Coca-Cola	
	Group A	Group B	Group C	Wajir [a]	Nairobi	Ghana	Nigeria	Cycle 1 [b]	Cycle 2
	Kenya			Kenya				Nigeria	
Treatment	44%	47%	51%	12%	37%	22%	27%	n/a	49%
Control	50%	50%	50%	11%	30%	20%	26%	n/a	55%

[a] In Wajir (Kenya), reported attrition rates between Round 1 and Round 2 are for Year 1 (first row) and for Year 2 (second row).
[b] For Cycle 1, baseline and endline data are not merged. No reliable data provided for attrition.

In both Coca-Cola's Cycle 1 and Cycle 2, the only reasons for attrition provided in the project's reports are migration within or outside the state and marriage. Furthermore, the research for the evaluation of Cycle 1 was carried out during the farming season, which might partially explain the dropout that occurred between baseline and endline. However, the project did not provide any data to explain the reasons for dropout.

Avanti's attrition rate was higher than what was envisaged at baseline. The main reasons given by the schools were that girls had either repeated a class or had transferred to other schools, with the latter being the most common. Girls that had repeated a class were not tested at endline because the difficulty of the tests was adapted for the cohort standard that was to be tested (Standard 6) and the project believed they would not be able to respond to the majority of the questions, leading to floor effects. High attrition rates in both treatment and control schools reduce the sample size and potentially limit the validity of comparing change from baseline to endline since the girls who left may have different characteristics to those who remained.

Avanti dealt with the loss of sample size by adding "substitute" girls to their samples for all of the intervention groups. This can also limit the validity of comparing change from baseline to endline if the substituted girls are different from the original sample. While checks for this found that re-contacted and substituted girls shared similar characteristics, there were some differences between the learning outcomes of the substituted girls and the girls from the original sample that were successfully re-contacted. In Group C, differences in literacy scores between the re-contacted group and the substitute group were statistically significant and in Group A, differences in literacy skills between the two groups was statistically significant.

Discovery provided very detailed data for attrition, and reasons for attrition (see Table 11). At baseline, the samples for all of the areas covered by Discovery were boosted by 30% to account for any attrition over the years. Furthermore, any bias due to attrition on balance in the composition of the cohort sample has been addressed through appropriate weighting of the endline sample.

In Wajir (Kenya), attrition was very low, but it is important to consider that the two data collection points (at Round 1 and Round 2) are only one year apart. With 11.4% at Round 2, cohort girls' attrition did not exceed this buffered attrition across the life of evaluation. For most of these girls, reasons for dropout were unknown or the girls were enrolled, but not present at school during Round 2.

In Ghana and Nigeria, attrition was below the 30% buffer, therefore there was no immediate threat to the sample size required at cohort level to capture project effects with the power that was calculated after baseline. In both countries, 40% of lost girls from both groups had transferred to other schools. In Nigeria, the main reason for these transfers was migration or relocation of families for jobs or the nomadic nature of their livelihoods. Head Teachers reasoned that severe illnesses, poverty, taking on employment to earn money for the households, parents not understanding the value of education, and early marriages led to 10% of dropouts among girls that were lost. In Ghana, reasons for attrition have not been detailed in the project's report, which leaves 50% of dropouts among lost girls unexplained.

In Nairobi (Kenya), cohort attrition of 34% was experienced at endline but that did not affect the power of the study largely due to the buffer. Attrition was higher in non-formal schools (44%) than in formal schools (27%) in both groups. The major reason provided by the Head Teachers for higher attrition of cohort girls in non-formal schools was parents' deliberate decisions to pull their children out when they reach Primary School Grade 6. Indeed, enrolling their children in either formal schools or rural schools elsewhere increases their chances of joining secondary school after sitting for their KCPE exams[38].

Table 11: Attrition rates and reasons for attrition in Discovery intervention areas

Attrition between baseline and endline for Discovery		Wajir [a]		Nairobi		Ghana		Nigeria	
		Year 1	Year 2	Treatment	Control	Treatment	Control	Treatment	Control
Number of girls[b] in sample	Baseline	361	340	1015	955	1071	907	1109	1021
	Endline	323	298	638	668	835	722	805	755
	Attrition rate	12%	11%	37%	30%	22%	20%	27%	26%
Reason for Attrition	Transfer	24%	36%	88%	91%	41%	39%	41%	39%
	Dropped-out	8%	5%	1%	1%	48%	56%	10%	24%
	Unknown reason	42%	26%	11%	8%	11%	5%	50%	37%
	Enrolled but not present	26%	33%	-	-	-	-	-	0%

[a] In Wajir, reported attrition rates between Round 1 and Round 2 are for Year 1 (first row) and for Year 2 (second row).
[b] Number of girls in the sample excluding substitutions. The attrition rate is the difference between the baseline and endline samples.

Traditionally, Uwezo is scored using a seven-point scale but **Avanti and Discovery projects used two different adaptations of Uwezo to report literacy and numeracy scores**, by changing the content and difficulty of questions and using non-standard scoring scales. These adaptations were necessary to ensure that the assessments were relevant to the context of the country and project and can be used to assess change from baseline to endline. However, the adaptations limit comparability between the two projects.

Inaccurate learning assessment scores

For the Coca-Cola Cycle 1 evaluation, the endline external evaluator raised concerns about the reliability of the baseline data for literacy collected by the previous external evaluator, and noted that the data might have been exaggerated, which led to an underestimation of the girls' improvements between baseline and endline. In parallel, it appeared that several administration errors for the endline literacy tests have been discovered during the EM data analysis. Indeed, as detailed in Annex D, Box D.2, the FM and the EM agreed that there must have been numerous administration errors in the reading fluency task that were not confirmed by the project. Nevertheless, the FM noted that as the administration errors are equally likely in both treatment and control groups, it is still possible to compare words per minute (oral reading fluency) scores between baseline and endline. Accordingly, the EM proceeded to this comparison but added some caveats that must be taken into account in the interpretation of the results.

Inability to match baseline and endline observations

Due to the reliability issues around Coca-Cola's Cycle 1 baseline data, the endline external evaluator did not re-contact the whole out-of-school cohort, and did not merge baseline and endline data for learning in order to conduct an impact evaluation.

---

[38] In Kenya, following the launching of the Free Basic Education Policy that established compulsory free basic education for all children, the public schools (also referred to as formal schools) became overpopulated resulting in overcrowded classrooms, insufficient learning materials and school furniture, a shortage of teachers, and poor quality of education. In addition, only the cost of attending schools is free of charge while the families are still responsible for covering the costs of auxiliary items such as uniforms, books, and activities. As for consequences, many parents choose to enrol their children in non-formal schools that are supported by communities, religious groups and other organisations, which are cheaper than formal schools. However, children attending these non-formal school have a lower access to public secondary schools due to a quota system introduced by the Ministry of Education in 2010 favouring children enrolled in formal schools. Consequently, some parents deliberately chose to take their children from non-formal schools to send them to formal schools so they can sit the Kenya Certificate of Primary Education (KCPE) examination in Grade 8 to increase their chances of being placed in a government secondary school.

Disparities between baseline and endline in the quality of Discovery's data limited the EM's ability to make comparisons between these data collection points

At baseline, it appeared that some of Discovery's enumerators were not initially properly trained in identifying best teaching practice, but training was improved for subsequent rounds of data collection at midline and endline. This resulted in the external evaluator disregarding the baseline data in the endline evaluation. As the EM, we chose to compare the data but add some caveats about the limited quality of baseline data.

Quality of control schools' registry data limits our ability to measure attendance

In Avanti's control schools, registers appeared not to be properly kept, which did not permit the project to present reliable data for attendance, and prevent any counterfactual analysis comparing the changes observed in the treatment group between baseline and endline to the control group.

In Discovery intervention areas – Nairobi (Kenya), Ghana and Nigeria – the external evaluator encountered several schools from both treatment and control groups, where the Head Teachers misplaced or lost registers from the previous school year due to several teachers transferring to another school and not returning the class registers. Moreover, collecting data on dropouts / transfers was challenging because of poor documentation and unavailable records. The approach used at baseline and endline was to ask the Head Teacher or Deputy Head Teacher about the number of students that had dropped out or transferred from each school, to review the registers for each grade (when available) to look for prolonged absenteeism, and to consult with the respective class teachers to confirm if students transferred or dropped out.

School-based sampling strategies limits our ability to measure enrolment or retention

Two SPW projects (Coca-Cola and Discovery) sampled girls in school, and then conducted household surveys of the sampled girls' households.[39] This means that the household survey is not representative of their larger community. Avanti randomly selected households in target areas but the survey questionnaire did not include questions about girls' enrolment and retention. As a result, it is not possible for SPW projects to calculate 'global' enrolment or retention rates in their target communities (the number of girls enrolled as a percentage of the total eligible girls in the community; or the number of girls enrolled in the previous year who have dropped out of school). Coca-Cola and Discovery surveyed out-of-school girls separately, which helps us understand the characteristics of these girls and their families. However, as mentioned above, they do not provide a representative sample to calculate enrolment and retention rates.

Potential contamination issues

In the three areas where Discovery sampled a treatment and a control group, teachers' transfers from treatment to the control school have been reported. These transfers were beyond the control of the project, and their effect on the project impact could not be measured. In Nairobi (Kenya) several schools were closely located to each other with some schools potentially indirectly affected by the project with the control group teacher and student obtaining information on teaching materials, and possible improvements in the community attitude on girls' education in both treatment and control groups, which favour learning. The project also mentioned that there were a few instances of potential partial compliance (for instance, some control school teachers showing up for training) but it was detected on time and rectified. For Avanti, some teachers and some parents mentioned they had transferred their child from control schools and some from intervention B schools where there were no ICT labs to schools with ICT labs so that they could get an opportunity to benefit from these resources. Some parents were reported to have transferred their children to schools that benefited from the stipend for similar reasons. In all the Discovery project areas, similar interventions and trainings were rolled out by other NGOs in treatment and control groups. This is further discussed in the effectiveness section.

These issues around contamination were outside the projects' control and could not be measured, but were an evident impediment to assessing the improvement in learning and attendance reported by projects. Therefore, we are not able to report on the extent to which contamination effects have played a role in the evolution of learning outcomes in treatment and control groups.

---

[39] This applies to Coca-Cola's group of in-school girls. While all interventions were delivered through the Safe Spaces set up by the programme, the girls were selected by choosing schools and selecting girls in the school. The out-of-school group was selected by finding girls who are out of school in the same communities.

3 Key Findings

3.1 To what extent has the SPW reached marginalised girls?

Key Findings

SPW projects aimed to reach marginalised girls. The projects' definitions of marginalisation did not change between baseline and endline and included aspects of socio-economic, educational and geographic marginalisation. Projects seemed to have an increased understanding of marginalisation in their project areas in the endline reports but have not adapted targeting or activities to reflect these changes.

According to the project endline reports SPW projects have reached 380,350 girls and 426,920 boys. Avanti estimated that 10,497 families and 115 merchants have been reached whereas Coca-Cola estimated that they reached about 40,000 community members. Discovery and Avanti considered all the girls in their project areas to be marginalised and delivered activities based on a whole school approach. There is limited information available from these projects at endline on whether they reached the most marginalised girls in their regions and communities, or which groups of girls make up the most marginalised in the project.

Coca-Cola was the only project that targeted specific girls for all of its project activities. Girls self-screened for participation based on criteria that included factors such as being an orphan, parents being sick, having a disability or being married before the age of 18. The EM reanalysis of the data suggests that they had some success in reaching the most marginalised, particularly orphan girls in single headed households.

Avanti's stipend component used a combination of targeting criteria based around school attendance and validation by school to target those perceived to be most at risk of dropping out of school. The project reports that 55% of stipend recipients improved their attendance but there remained unmet financial barriers to school attendance and factors such as 'sickness' which account for a large number of girls who missed school.

There is limited information available about the reach and visibility of the SPW projects in the broader community. EM reanalysis of the data showed that awareness of organisations working on girls' education was highest in Coca-Cola project areas and many caregivers reported that girls had received tutoring activities in Discovery and Coca-Cola areas. The picture on reach into the broader community is less clear. For example, EM reanalysis data showed that very few people in the project communities have seen Discovery's TV show.

3.1.1 How are SPW projects targeted?

Marginalisation

The GEC aims to "expand education opportunities to marginalised girls."[40] The 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."[41]

DFID did not prescribe the factors that marginalised girls from education as there was a lack of evidence about what caused marginalisation, and these factors were likely to vary from one context to another. GEC applicants were encouraged to focus on the girls with the greatest education needs in their target communities, leaving it to them to explain how and why the target girls were marginalised from education. The SPW tended to take similar approaches to defining marginalisation in terms of a combination of socio-economic, geographic and/or educational criteria.

Three broad categories of marginalisation were identified across the SPW projects at baseline. These categories did not change between baseline and endline.

- **Educationally marginalised girls:** All projects used educational outcomes to define marginalisation at baseline, with the focus being on out-of-school girls, or girls from households that had completed very few

---

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

years of education. In endline reports projects talked about additional criteria to define marginalisation which included poor attendance, low levels of enrolment, low motivation to stay in school, poor classroom participation as well as low learning outcomes and being seen as at risk of dropping out.

- **Socio-economically marginalised girls:** Projects utilised a number of different criteria in their definitions of socio-economic marginalisation. In the projects' baseline and endline reports these included: girls whose families are unable to meet basic needs or pay school fees, orphan girls, girls with disabilities, girls facing early marriage or a young pregnancy, and girls who have a sick parent or caregiver. More broadly the socio-economic barriers to education included any other definitions that fit the context, such as belonging to nomadic or pastoralist communities.

- **Geographically marginalised girls:** SPW partners identified geographic areas where they considered girls to be generally marginalised. The criteria used to select these areas included: the presence of widespread household poverty, low school enrolment and learning outcomes, and negative attitudes towards girls' education.

  The map in Section 1.1.3, Figure 1 shows the location of the areas that each project worked in. Discovery worked in Nairobi and Wajir counties in Kenya, the Northern region of Ghana, and Kano State in Nigeria. Coca-Cola worked in four Northern Nigerian States (Kano, Kaduna, FCT and Lagos), and Avanti ran activities in four counties in Kenya (Kilifi, Kajiado, Makueni and Uasin)

At baseline, projects were able to select the marginalised groups they wanted to target through their interventions. Discovery and Avanti took the view that all the girls within their target area were marginalised and so did not target the majority of their activities towards a particular demographic within that broad criteria. In their endline report Avanti described girls in the project areas as exposed to multiple marginalisation factors in education, economic and social spheres. Avanti further stated that the project was able to reach educationally marginalised girls who have poor performance or attendance or are at risk of dropping-out of school, by providing tutoring and other learning materials that enabled them to catch up. Discovery described all the areas where they delivered activities as "where poverty is highly prevalent and the population is transient" in their project endline reports. As such, all of the girls in the area were considered to be marginalised by being out of school or at risk of dropping out of school. Discovery aimed to reduced marginalisation through improving school enrolment, attendance and girls' motivation to stay in school.

Some qualitative data in the endline reports described how some of the marginalisation factors identified impact on girls' lives in the target communities. For example, early marriage and pregnancy, family migration, language or girls needing to take up employment to support their families. In some cases, the projects identified the same issues at baseline – for example, girls with household responsibilities had to complete them after school with this often taking priority over schoolwork. The endline reports also mentioned a number of additional external factors that may have increased the level of marginalisation experienced by adolescent girls over the project period and which were not accounted for in the project design. For example, in their endline report Coca-Cola Cycle 2 mentioned heightened conflict and restrictive social and religious norms around girls' education (particularly resulting from the Boko Haram attacks) which have increased marginalisation.

Individual Target Groups

Table 12 below shows the target groups of the different projects. All three projects targeted in-school girls, the exception being Coca-Cola which targeted in-school and out-of-school girls with different activities. Two out of the three projects considered in this report (Avanti and Discovery) targeted students in primary school. Since their activities worked at either the classroom level or the school level, interventions were not targeted at the individual level within schools. Discovery worked specifically with Grade 4 classes in all schools, whereas Avanti worked with all primary grade levels. Because a significant portion of students have repeated grades due to missed school time, this does not map directly to the ages of girls targeted. Table 13 shows the ages of girls in each of Discovery and Avanti's project areas at baseline, based on data from their survey of girls in school. If all students had started at the correct age and did not repeat grades, for all project areas in Discovery they would be 9 years old. In Kenya, primary school lasts through grade level 8, so girls aged 6-13 are primary-age, whereas in the other countries, primary school goes through level 6, so girls aged 6-11 are primary age. From Table 13, we can see that the ages reached through the grade-level targeted varied across project areas. In the Ghana project area in particular, a large portion of the beneficiaries were secondary-age girls even through the intervention targeted primary schools.

Coca-Cola was the only partner to target girls of a certain age – 16 to 19 – which corresponds to secondary school aged girls.

Table 12: Who is targeted by SPW projects?

	Avanti	**Discovery**	**Coca-Cola**
Individual targeting	X *(with exception of stipends)*	X	✔
In-School Girls	✔	✔	✔
Out-of-School Girls	X	X	✔
Primary school girls	✔	✔	X
Secondary school girls	X	X	✔

Table 13: Ages of Girls in Discovery and Avanti Project Areas (Source: Project Survey of Girls)

	Avanti	**Discovery**		
	Kenya	**Kenya**	**Ghana**	**Nigeria**
5 to 8	40%	10%	3%	14%
9 to 11	44%	83%	44%	63%
12 to 13	12%	5%	38%	20%
14 to 15	3%	0%	14%	3%
16 or older	0%	1%	0%	1%

Socio-economic groups: The only project which targeted all their activities at the individual level based on criteria of social and economic marginalisation was Coca-Cola. Girls needed be aged between 16 and 19 years at baseline and meet at least one marginalisation criterion. The criteria used were slightly modified between the first and second cycle of the project. The original criteria were:

- Girls married before 18
- Girls who are pregnant or had a child or children before the age of 18
- Girls who are divorced or widowed
- Girls who have a disability
- Girls who are an orphan or come from a single headed household
- Girls who come from a household with a sick parent or husband
- Girls whose parents are not able to pay school fees

After the first cycle, the last marginalisation criterion was removed.

None of the SPW projects have changed their conceptualisations of marginalisation or targeting since baseline. Projects were aware of social and economic marginalisation experienced by girls in their project areas and there is some evidence in the endline reports that the projects have gained a greater understanding of the dynamics of girls' marginalisation within their target communities between baseline and endline. For example, the increased attention to the languages spoken and to what degree of fluency between the baseline and endline reports from the Coca-Cola projects demonstrates this. At baseline the main languages spoken are given for in-school girls but for Cycle 2 endline the evaluators give more detail about how many girls speak each language (Hausa, Arabic, English, Igbo, Yoruba, other). This shows that 70% of girls in Federal Capital Territory speak English compared to 8.4% in Kano and 18.6% in Kaduna. The low levels of English proficiency are likely to have impacted on girls' ability to learn in the classroom and their scores in EGRA and EGMA tests. The Cycle 1 endline report found higher levels of English spoken in Kano, Kaduna and the Federal Capital Territory compared to the reports for Cycle 2, however the evaluators noted that teachers were translating the subject matter to teach to the girls in their first language which was most likely to be Hausa.

Avanti targeted the stipend element of their programme. The criteria for inclusion were primarily based around school attendance, as those with low school attendance are assumed to be at risk of dropping out of school. The assumption was that girls were not attending school because their families could not afford schooling. According to its endline report, Avanti identified students to receive the stipend based on the following criteria:

- Girls whose attendance was less than 70% according to the school monitoring system (comprising 80% of the targeted beneficiaries)

- Boys whose attendance was less than 50% according to the school monitoring system (comprising 10% of the targeted beneficiaries)

- Students nominated by the school to ensure school cooperation (comprising 10% of the targeted beneficiaries)

- Exclusion of any selected student who had been selected on the basis of attendance to receive the stipend where schools and school committees advised that they were not in need compared to other students.

The process was tweaked according to the school and the final selection of stipend beneficiaries was validated by school committees of teachers and senior teachers. It is not clear how well the final beneficiaries met the criteria, but the process demonstrates how areas of marginalisation can be combined to target the marginalised and strengthen the school and community's sense of inclusion and ownership.

Presence of project individual target groups for targeted projects/ interventions

As part of their baseline research, projects collected data on the presence of their planned target groups as well as a control group of marginalised girls who did not receive the intervention but had similar characteristics. The expected target groups were largely found to be present with the exception of Coca-Cola, who did not find disabled girls or widowed/ divorced girls within the communities they targeted. At endline, there was limited information available about the presence of the target groups in the community areas and whether the project had effectively reached these groups. The available information is summarised in Table 14.

Table 14: Project targeting - primary target groups by project

Endline Project Targeting	Coca-Cola Cycle 1	Coca-Cola Cycle 2	Avanti
	Nigeria	Nigeria	Kenya (stipends component)
School Phase			
Lower Primary			✔
Upper Primary			✔
Secondary	✔	✔	
Socioeconomic Group			
Disabled Girls	✦	✔	
Girl with sick parent	✔	✔	
Orphan/one parent	✔	✔	
Poor	✔	✔	⊘
Rural			
Slum/Periurban			
Married Girls	✔	✔	
Young Mothers	✔	✦	
Minority/Migrants			
Widowed/Divorced (child)	✦	✔	
Socially marginalised			⊘
Other			
Educational Group			
Go to informal schools	✔	✔	
Out-of-school	✔	✔	
In school girls	✔	✔	✔

Endline Project Targeting	Coca-Cola Cycle 1	Coca-Cola Cycle 2	Avanti
	Nigeria	Nigeria	Kenya (stipends component)
Poor school attendance			✓

Key: ✓ target group reached, ✦ target group not reached or not present. ◌= no evidence of whether the target group was reached. Target groups not included in the project design are shaded grey.

## Effectiveness of targeting at the level of individuals or communities

One important question about project reach is whether targeting girls at the individual level is more effective at reaching a marginalised group than targeting only at the level of communities. As Coca-Cola targeted at the individual level and Discovery and Avanti targeted primarily at the community level, the SPW offers an ideal opportunity to compare these approaches. Coca-Cola selected girls to be included in the project by administering a short screening questionnaire to determine if the girls met the marginalisation criteria. The results of that survey for in-school girls included in the project are presented in Figure 5 and Figure 6 respectively.

Figure 5: Coca-Cola, Cycle 1 - Marginalisation screening results at baseline (source: project's household survey)
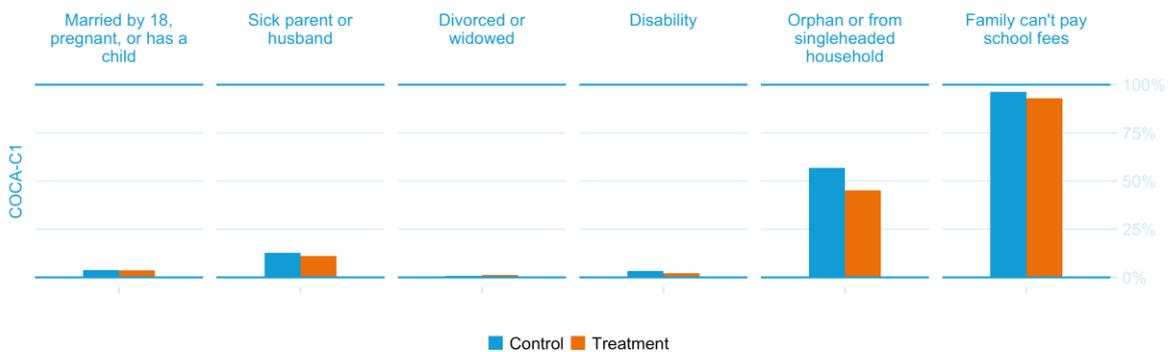
In Cycle 1, the most common area of marginalisation was that the girls' parents were unable to pay school fees. This is intriguing because these are the screening results for girls who were currently enrolled in school. It appears that either the girls or the people administering the screening survey interpreted this criterion to mean that the family found paying school fees to be a financial burden. However, there is no way to know what criterion was used to make that assessment and since nearly all girls were reported to meet this criterion, it appears to have undermined the goal of screening for girls who are part of particularly marginalised socioeconomic groups. This screening criterion was not used in Cycle 2, most likely due to this problem. However, it is worth examining the evidence of whether this rough screening mechanism succeeded in reaching the poorest members of the communities.

Figure 6: Coca-Cola, Cycle 2 - Marginalisation screening results at baseline (source: project's household survey)
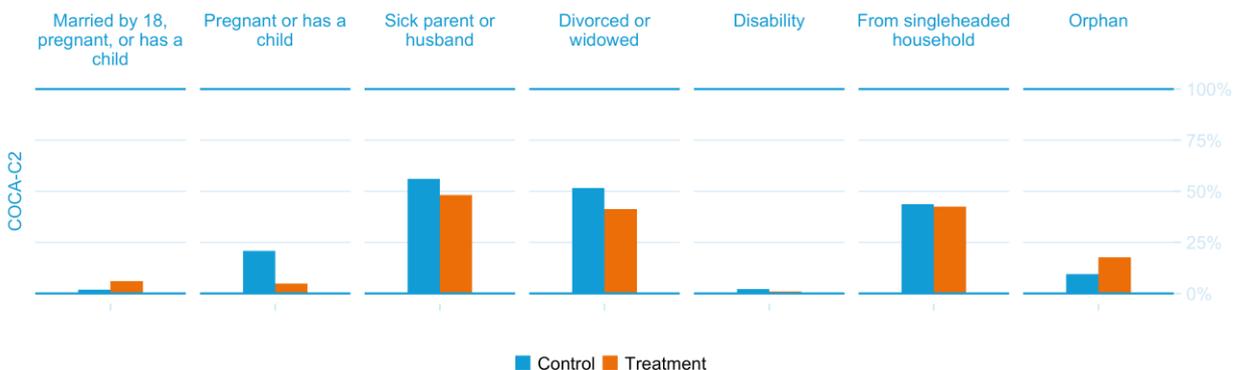
First, taking the screening criterion at face value, we compare the portion of families that reported having to pay school fees across the GEC. This comparison is presented in Figure 7. We don't have a directly comparable

question for Avanti, but in a survey question that asked the amount of money they paid to the school, all households reported some positive amount. It is hard to tell for sure if that question was correctly interpreted by households, but it is broadly consistent with the picture presented in Figure 7, which shows that households in Kenya are the most likely to have to pay school fees. This evidence suggests that the families in Coca-Cola Cycle 1 did not face the largest burden across the SPW in having to pay school fees. Moreover, in Coca-Cola Cycle 2, this screening criterion was dropped and the portion of parents who reported having to pay tuition remained at a similar level. This evidence suggests that this screening mechanism did not effectively select a group for whom school fees were a particularly strong barrier. The somewhat higher proportion of families that had to pay school fees in Coca-Cola's project compared to Discovery Nigeria is likely explained by the fact that primary schools are less likely to charge fees than secondary.

Figure 7: Discovery and Coca-Cola - Portion of families who have to pay school fees according to primary caregiver (<u>source</u>: projects' household surveys)
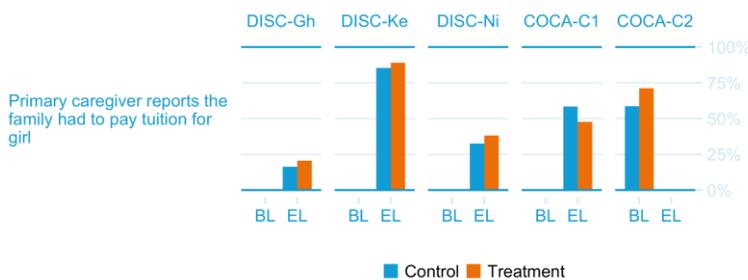
On the other hand, if we assume instead that the tuition screening criterion was only meant to be a proxy for how poor households felt, there is some evidence that Coca-Cola Cycle 1 did capture a group of households that felt poorer than other project areas. In Figure 8, we present the responses to a household survey question asking the primary caregiver how they feel about their family's financial situation. It is important to note that these results come from the household survey which was conducted after the girl was already selected to be in the sample, so there would not have been pressure for the households to exaggerate their need to be included in the programme. At baseline, a much larger portion of households in Coca-Cola Cycle 1 reported feeling very poor compared to Cycle 2 in all the Discovery project areas, including Discovery's Nigeria project area. This provides an initial indication that even this rough screening method might have helped target a group that at least felt they were poorer than average. However, this should be interpreted with some caution since the proportion of families reporting feeling unable to meet basic needs declined almost by half by endline (9 months later), which is more in line with the other projects. Since there is no particular evidence of a widespread reduction in poverty during this period, it is possible that the baseline result was simply the result of inconsistent data collection.

Figure 8: Discovery and Coca-Cola - Experience of poverty as reported by primary caregiver (<u>source</u>: projects' household surveys)
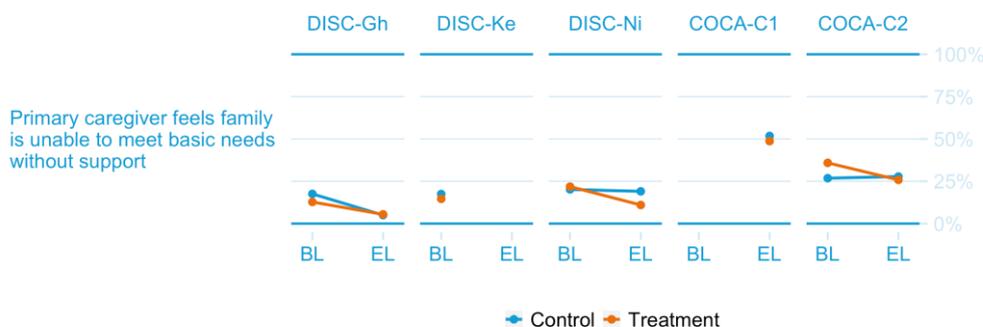
Avanti provides another interesting point of comparison. To select its sample of schools the project was supplied with statistical data on schools for selected counties by the Kenyan Ministry of Education. From this data, Avanti selected its school sample based on size, access to electricity and participation in the government's national sanitary towel programme (the main indicator for marginalisation used by the project). Using access to a reliable source of electricity as a key selection criterion may have inadvertently created some bias toward relatively wealthier school communities. On the other hand, in their stipend programme, Avanti attempted to target the

neediest girls on an individual level. In Figure 9, we show the income distribution of targeted families from baseline to endline. Income is given in unadjusted local currency, but the orange line in each histogram shows the international poverty line of US$1.9 per day (in purchase-parity adjusted 2011 US dollars). It is somewhat difficult to compare this income distribution to the national average since the last poverty headcount comes from 2005, when 43.4% of the population lived below the poverty line. Compared to that benchmark, in Avanti's project areas 38.4% were below the poverty line at baseline and 41.6% at endline. If we account for progress since 2005, this is probably close to the national picture. However, it is notable that there is a spike at the top of the income distribution[42] at baseline. We can see from these income distributions that even in areas that are not worse off compared to the national average, more than one third of the population meets the international standard of extreme poverty. This is a good illustration of the fact that as more and more countries enter middle income status, a strategy to reach the poorest will have to target not just uniformly poor communities, but also unequal areas. Unfortunately, we do not have data on how well Avanti's stipend targeting worked at reaching the poorest members of these communities.

Figure 9: Avanti - Histograms of income distribution of project families at baseline and endline compared to international poverty line (source: project's household survey)
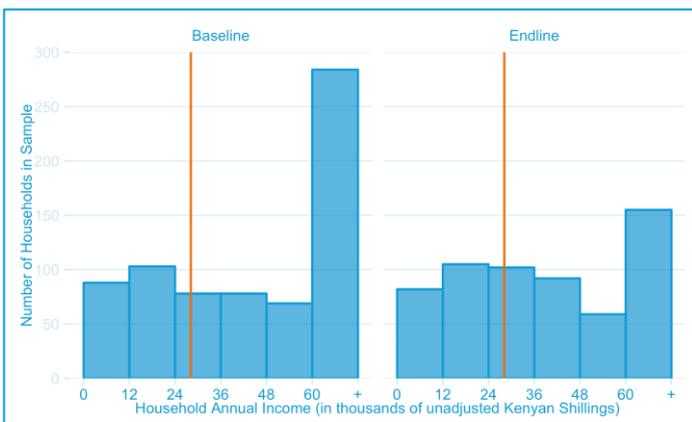
Returning to the results of Coca-Cola's screening process, the other major group reached through their screening process was orphans and girls from single-headed households. This was consistent between both cycles, and in interviews with the EM Coca-Cola staff confirmed that orphans had been the main focus of their targeting of in-school girls. In this area it is clear that Coca-Cola's screening process succeeded in targeting more girls in this subgroup than projects that only targeted at the community level. In Figure 10, we show a comparison of both cycles of Coca-Cola to Discovery in terms of the number of girls who meet this criteria. While there are some minor differences from the trends found in the screening survey given directly to girls, this confirms the overall pattern that Coca-Cola was able to reach a significant proportion of girls in this subgroup.

---

[42] Data was collected in income ranges, so the highest income category was 60,000 Kenyan Shillings or more, therefore we don't know exactly what the upper end of the income distribution looks like. In purchasing parity adjusted US dollars, 60,000 Kenyan Shillings is equivalent to approximately 1,300 USD per year.

Figure 10: Discovery and Coca-Cola - Portion of girls from single-headed household or orphans (<u>source</u>: projects' household surveys)
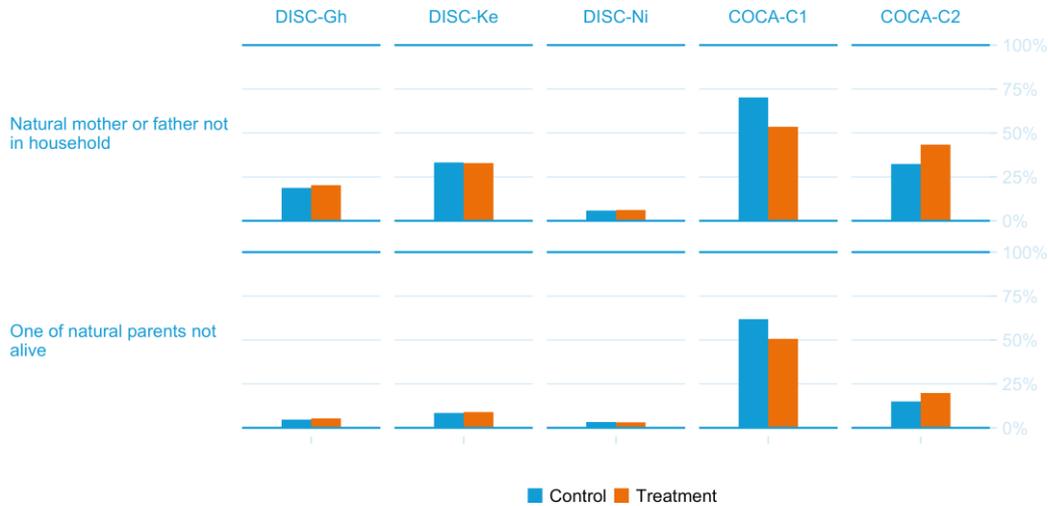
Finally, we considered disabled girls. Coca-Cola attempted to reach this group, but according to the screening results, reached a very small number of girls in this subgroup. In Figure 11, we show the proportions of girls with disabilities across Coca-Cola and Discovery's project areas, using data from their household surveys. It is interesting that whilst almost no girls were reported to have disabilities in either Coca-Cola Cycle 1 or 2, the household survey shows that nearly 25% of girls in Coca-Cola Cycle 2 had a disability. Given some of the reliability problems with Coca-Cola's data, this could simply be an error. However, this could also be an indication that it is difficult to identify girls with disabilities because they do not recognise their disability or may be sensitive about discussing it. The household survey contains a series of questions about specific disabilities ranging from vision to communication, and so parents may be willing to report disabilities that the girls don't consider significant. Also notable in these results is the high proportion of disabled girls in Discovery's Kenya sample. Discovery's Kenya sample also had a higher proportion of girls from single-headed households, almost on par with Coca-Cola Cycle 2. Taken together, this suggests that in some cases selecting the neediest communities can also be a very effective way of reaching certain marginalised groups.

Figure 11: Discovery and Coca-Cola - Proportion of girls with disability (<u>source</u>: projects' household surveys)
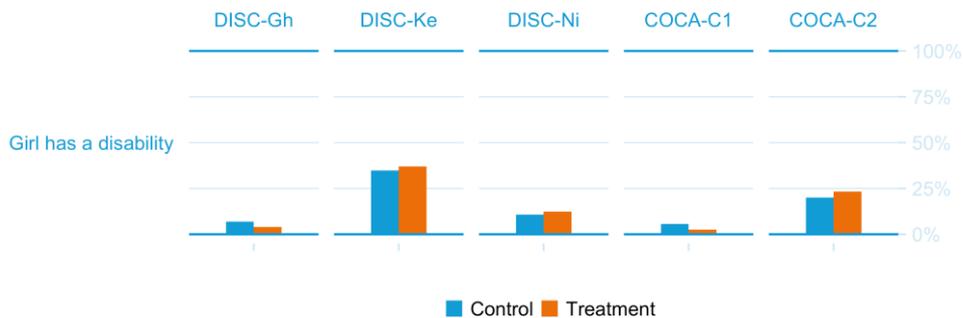
3.1.2 Who did the Projects Reach?

Number of beneficiaries and awareness of activities

Overall there was very limited data available in the projects' endline reports on how wide the reach of each project was, whether the projects were able to reach into the broader community or the most marginalised girls. Table 15 shows monitoring data on the number of girls reached with the project activities. Overall the SPW projects have reached 380,350 girls. It is important to note that information presented in Figure 11 and in Table 15 comes from two different sources of data: the percentages in Figure 11 come from the data-reanalysis which corresponds to

what the parents reported, whereas numbers displayed in Table 15 come from the projects' reports. Discovery did not count the number of children with disabilities because it was not set as a criteria for targeting the beneficiaries.

Table 15: Number of girl beneficiaries reached

	Avanti	Discovery	Coca-Cola	Total
In-school	56,561	302,627	7,128	367,250
Out-of-school			14,034	13,154
Total	56,561	302, 627	21,162	**380, 350**
Subcategory				
Disabilities	30	N/A	2,509	2,539

In the endline reports, projects reported on their estimated reach beyond targeted girls - Table 16 summarises how each project described these beneficiaries. For example, in Coca-Cola's endline report they estimate that 40,000 additional girls would have benefited from the project due to mentoring and relationships with girls who have participated in the project activities and would benefit from effects such as attitude changes. The whole school approach taken by Avanti and Discovery means that other children in the schools are bound to benefit from activities.

**Table 16: additional beneficiaries of the project**

	Avanti	Discovery	Coca-Cola
Primary Target group	In school girls	In school girls	In school girls Out-of-school girls who fit the project definition of marginalization; Girls with children, married, divorced or widowed girls, orphans, girls with a sick parent in the household
Other beneficiaries	In school boys: number of boys in class who are exposed to project activities equally to girls - **99,073** Teachers: number who benefit from training or related interventions – **4,613** Community members: number reached through community messaging, meetings, advocacy and families who benefited from the stipend – **10,497 families. 115 merchants**	In school boys: number of boys enrolled in schools participating in the project – **327,847** Teachers: number trained as part of intensive teacher training – **15,694** Community members: not tracked.	Other female students: number of girls who benefit from the interventions in a less direct way, for example through being mentored by girls who participated in the project – **estimated at 40,000 girls** Teachers: number who benefit from training or related interventions – **estimated at 750** Community members: number reached through community messaging, meetings, advocacy and participants in community groups – **estimated at 40,000**

There is limited data available on the visibility of projects within the community. Avanti and Coca-Cola described the number of community members they have reached in their endline reports, however Discovery did not track the number of community members they reached with activities to build support for learning centres.

The project endline reports provided some qualitative evidence that parents have become more supportive and encouraging towards girls' education. Coca-Cola estimated that they have reached 40,000 community members, including parents, community leaders and school management committees through community events, meetings and advocacy visits. Despite this qualitative evidence, the endline project report showed that girls feel they are at risk of not being able to complete their education due to factors including the unsupportive attitudes of family members to education, and particularly being forced into marriage.

Awareness of the projects within communities appears to be quite low. One of the questions in the household survey asks if organisations worked in their community last year to improve education. As Figure 12 shows, the reanalysis of responses here shows no difference between treatment and control areas, as has been the case across the GEC windows.

Figure 12: Discovery and Coca-Cola - Perceptions of interventions to improve education (source: projects' household surveys)[43]
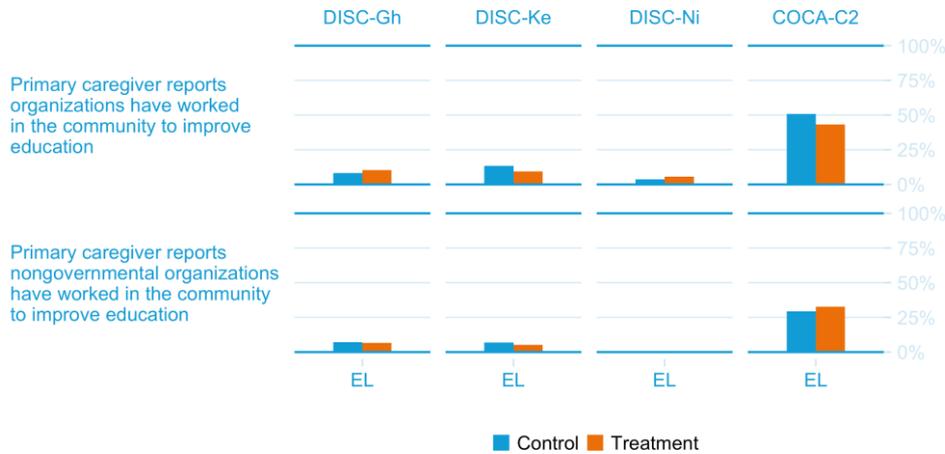
There are a number of potential explanations for this, for example the lack of difference between treatment and control groups could reflect contamination between the groups, or that many actors are working in the area and the SPW makes up a small part of the overall effort. Respondents may also not think the projects are relevant when answering the question or may not be aware of them, particularly where branding has not been effective, or as Discovery staff suggest in interviews, the majority of activities have taken place within schools with no materials for students to take home or into the community. Interviews with project staff suggest that the greater access to education services in the urban project areas and geographical proximity of some treatment and control groups means that contamination of the project treatment area may have occurred.

When asked about the provision of direct services for girls in the community the EM reanalysis shows the clearest impact for Coca-Cola Cycle 1 (Figure 13). 100% of caregivers in the treatment group reported that the girls had attended special classes or study groups. However, there was no difference between treatment and control in the Cycle 2 areas, despite this kind of activity being closely aligned with the project activity. In interviews, the Coca-Cola staff suggested that there were limited actors working on girls' education in the project area but there are projects more focused on security which provide safe spaces. This kind of intervention could potentially have been perceived as study groups or support with schooling by caregivers.

Figure 13: Discovery and Coca-Cola - Perceptions of activities targeting girls directly (source: projects' household surveys)[44]

---
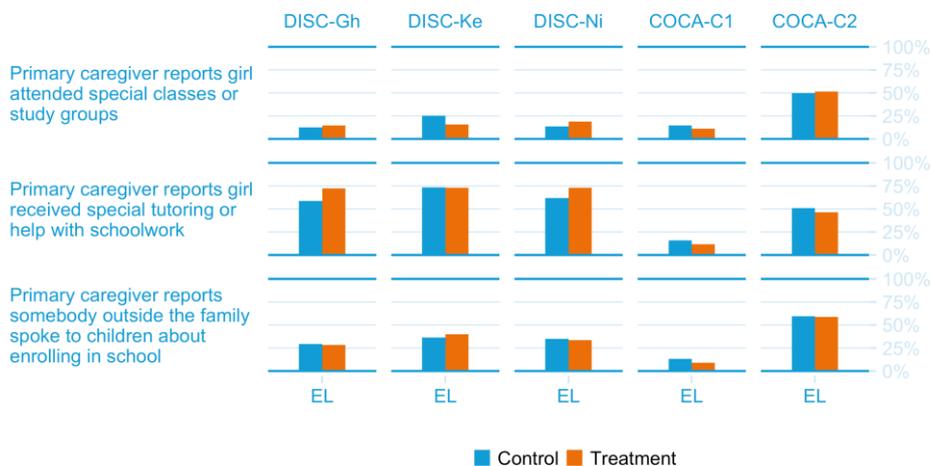
[43] These questions were not included in the baseline survey so we cannot look at change in these variables.
[44] These questions were not included in the baseline household surveys, so we cannot look at change from baseline.

These variables are not directly relevant to the activities delivered by Discovery, but the EM data reanalysis shows a large number of caregivers reported that girls are receiving special tutoring or help with homework. It is possible that the life skills based curriculum of Discovery's girls' clubs is being perceived as directly supporting education within the community, particularly as they take place on school premises and are delivered by teachers. But it is unlikely that this is a major portion of the tutoring activities that parents are reporting because there was little difference between treatment and control on this variable. This suggests that there may be other actors working in these communities providing a significant level of support to girls.

Lessons learned

- **Using a school-based selection approach to select their beneficiaries constrained Avanti's and Discovery's capacity to reach marginalised girls**. Overall, more boys have been reached by SPW projects than girls which potentially raises questions about the projects' contributions to reducing gender disparities. If school-based selection is used then targeting criteria are needed to at least identify which girls are more marginalised than other girls and boys to inform intervention designs that specifically meet their learning needs and address education inequalities.

- **Using marginalisation checklists to identify their target population** enabled Coca-Cola to define marginalisation as a relative concept and target girls who were more marginalised than others, by not assuming that every girl in a community is marginalised in the same way.

3.2 What impact has the SPW had on marginalised girls' learning?

Key Findings

There is little evidence that SPW projects had an impact on learning outcomes. Impact was assessed by comparing the change in learning outcomes from baseline to endline between the treatment and control groups. This difference was not statistically significant for most learning outcomes in most project areas. Of 13 tests of the effect of a project on a learning outcome, the difference between treatment and control in 10 was not statistically significant. The difference was only positive and statistically significant for literacy in Discovery Nigeria and it was negative and statistically significant for numeracy in Discovery Kenya and Ghana. This is consistent with a window-wide conclusion that there was no impact on learning outcomes.

Avanti did not have a statistically significant impact on learning outcomes, but the full package of treatments appeared to perform better than treatment combinations that excluded some elements. Results were better for the group that received all components of the treatment compared to those that did not receive personal tutoring or stipends. Tutoring appeared to improve literacy for high achieving girls and improve numeracy for low achieving girls. Stipends appeared to have the opposite effect.

Discovery had a positive impact in Nigeria where literacy skills were initially very low, but negative impact in Nairobi (Kenya) and Ghana, where numeracy skills were initially relatively better.

Coca-Cola did not have any statistically significant impact on in-school girls' learning. By contrast, some improvements in life and business skills have been observed among the out-of-school girl group.

Learning is one of the GEC's key outcomes. In this report, the term "learning" is used to describe girls' progress in school and the acquisition of new skills and knowledge in relatively broad terms. However, when measuring learning as a GEC outcome, we apply a more specific definition of learning as "a change in ability over time" in literacy (i.e. oral reading fluency), and numeracy skills.

Learning tests used across SPW

Literacy and numeracy were measured by using adapted versions of two international learning assessments, the Early Grade Reading Assessment (EGRA), the Early Grade Math Assessment (EGMA), and Uwezo reading and numeracy tests. Similar to baseline, the differences between these tests did not limit our ability to provide comparisons across projects.

- **Avanti** used an adapted Uwezo test. The literacy test included reading fluency tasks, which allow us to measure the number of words per minute (wpm) read by the girls. The paragraph used for this task had 100 words at baseline. It had been extended to 179 words at endline to avoid ceiling effects. For the same

reasons, numeracy test scores had also been extended from 10 to 32 points. To ensure comparability between baseline and endline, the project standardised the wpm and the numeracy scores, and conducted statistical tests on these new scores.

- **Discovery** used an adapted Uwezo test rated on 40 points both for literacy and numeracy. The literacy test did not include any oral reading fluency test, which meant that words per minute were not measured.
- **Coca-Cola** opted for an EGRA test for literacy, measuring the number of words per minute that the girls were able to read. For numeracy, it used an EGMA test scaled on 100 points. Only in-school girls were tested for improvements in literacy and numeracy. For out-of-school girls, the project assessed improvements in their financial literacy as per the theory of change.

*In both Cycle 1 and Cycle 2, only the in-school cohort was tested for improvements in literacy and numeracy, whereas the out-of-school girls were not tested and did not contribute to Coca-Cola's learning targets. Instead, Coca Cola's theory of change was designed to deliver **improvements in financial literacy**. The percentage of girls improving in financial literacy is based on a perception test and is equal to 51% and 36% of the girls that finished the programme in Cycle 1 and 2 respectively.*

Table 17: Learning tests and units used at baseline and endline across SPW

SPW - Learning assessments		SPW projects by country and region		
		Avanti	Discovery	Coca-Cola
Literacy				
Test		Adapted Uwezo (English)	Adapted Uwezo (English)	EGRA (English)
Unit	Baseline	wpm (max 100)	/40	wpm
	Endline	wpm (max 179)		
Numeracy				
Test		Uwezo	Uwezo	EGMA
Unit	Baseline	/10	/40	/100
	Endline	/32		

Benchmarks for literacy and numeracy and comparability of EGRA results across languages

International education experts consider oral reading fluency a strong predictor of later literacy[45]. Children who do not acquire basic reading skills at an early age are more likely to repeat grades and eventually drop out of school, whilst the performance gap between early readers and non-readers increases over time. In EGRA tests, it is generally assumed that students should be able to read **a minimum of 45-60 words per minute** in order to understand a simple passage of text (Figure 14).

Among SPW projects, only Coca-Cola used EGRA to measure literacy. By contrast, Avanti used an adapted Uwezo scale that included an oral reading fluency task which measured wpm, enabling us to compare the results to Abadzi's international benchmark.

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

---

[45] 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.

Figure 14: International benchmarks of oral reading fluency by grade and age (Abadzi, 2011)

Grade	Age	Expected reading fluency (wpm)	Year-on-year improvement (wpm)
1	6 years	21	21
2	7 years	45	24
3	8 years	63	18
4	9 years	85	22
5	10 years	90	5
6	11 years	108	18
7	12 years	110	2
8	13 years	110	0
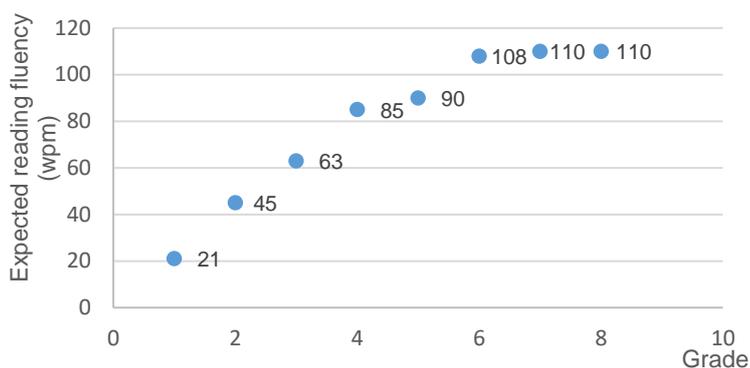
Tests not measuring words per minute

The Uwezo test assesses literacy by using levels rather than a score, questions or subtasks. It measures children's ability to perform literacy and numeracy tasks at a level of difficulty that is typical for Primary Grade 2 assignments. Ability is then reported as the level of tasks that the child can perform comfortably. Although this test allows us to report on students' learning progress, no benchmark on the level attained at each grade has been found.

Because both Discovery and Avanti used significant adaptations of the Uwezo test, neither partner reported their results on this standard scale: the former reported scores on 40 points for literacy and numeracy at both baseline and endline, and the latter on a 20 point scale at baseline, and 40 points scale at endline for literacy, and on a 10 points scale at baseline and 32 points scales at endline for numeracy. Each test covers slightly different competencies and does not allow comparisons between these tests and the standard seven Uwezo assessment levels based on Grade 2 competencies. Although a conversion of these raw scores to the same seven-level scale of competencies aligned with the standard Uwezo scale would be feasible through a subsection analysis, we chose to report the scores out of 40 due to the higher difficulty of Discovery's assessment compared to a standard Uwezo test. Avanti did not use standard Uwezo sub-sections, therefore we also chose to keep the original scoring. For literacy, we used the words per minute score measured in one of the sub-tasks of Avanti's adapted Uwezo test (that is not included in the standard Uwezo assessment).

Projects' achievements against SD targets

In addition to presenting projects' literacy and numeracy scores achieved at endline, we also indicated **whether projects have achieved their targets** as defined at baseline and expressed in standard deviations (SD). Projects' targets are the same for both literacy and numeracy, and were set by the FM and projects using baseline learning data[46]. By choosing to define the learning target as a proportion of the standard deviation of learning scores for girls, the GEC aims to account for the specific learning distributions among target populations, and hence to reflect the context in which each project intervenes.

At baseline, each project was asked to (see Figure 15):

- **Measure the learning scores** for each grade that composes the cohort group of girls enrolled, in the *grade in which girls from the cohort group would theoretically be at SPW endline collection point*.

- **Compute the standard deviation for each of these groups (SD).**

- **Set the learning target at 0.2 standard deviation (SD)** of scores for the relevant grade[47] (or 0.4 for Avanti).

- **Multiply the SD** of the relevant grade (computed in point 2.) **by the target** fixed (in point 3.) in order to obtain the **target score**.

---

[46] Fund Manager for the GEC (September 2013), *GEC Learning guidance on outcome targets*.
[47] 0.2 and 0.3 SD are commonly used as a target in educational programmes. An increase in learning scores of 0.4 SD and over is considered a very high achievement across educational studies. Source: Evans, David K., and Anna Popova. 2016. "What really works to improve learning in developing countries? An analysis of divergent findings in systematic reviews," World Bank Research Observer 31(2): 242-270.

At endline, **the percentage of achievement against the target** is obtained by dividing the impact measured by the difference-in-difference estimator (beta) by this target score.

Figure 15: Learning target setting for GEC projects[48]
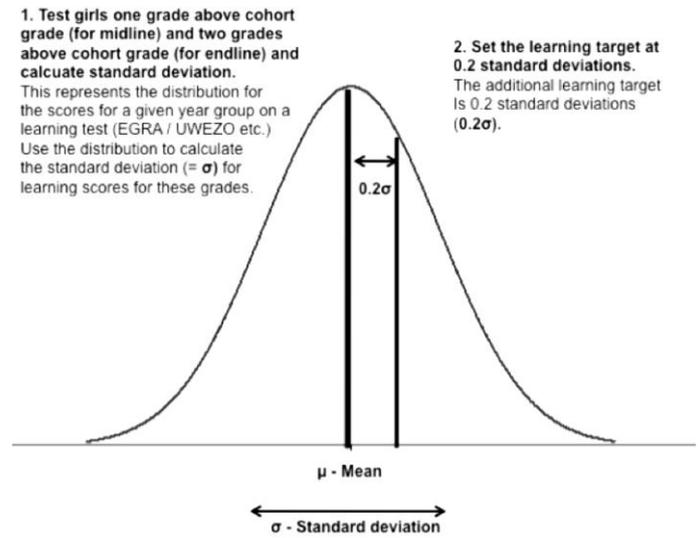
This process allows targets to be set that are specific to the projects' context. It also enables a **comparison of projects' achievement across different learning tests** (EGRA /EGMA, Uwezo, ASER, national assessments), since standard deviation is dependent upon the type of assessment used. In this report, as a rule of thumb, we consider that an impact in SD is negligible if it stands at less than +/- 0.05 SD for Discovery and Coca-Cola, and +/- 0.1 SD for Avanti, which corresponds to an achievement against target that is smaller than 25% if the impact in SD is positive, or greater than -25% if the impact in SD is negative.

With regards to Discovery's use of a Performance Evaluation (PE) approach in Wajir, where there is no treatment and control group, the percentage of achievement against target is found by replacing the difference-in-difference score by the change observed in the cohort between the two data collections points (i.e. a simple difference calculation).

At endline, both the EM and FM found that **none of the projects achieved their targets for literacy or numeracy**. The best achievement found was for literacy in Discovery (Nigeria) which achieved 71% of its target. It was also the only project location to achieve a statistically significant impact over and above the control group.

Coca-Cola Cycle 1 also achieved half of its target in literacy, but no difference-in-difference adjusted estimator is presented here (see Table 18). The reason is that for Coca-Cola Cycle 1, the baseline and endline groups were not matched. Therefore, the achieved result of this target has been derived by considering the difference between treatment and control groups, and between baseline and endline samples, among *different* girls.

---

[48] Fund Manager for the GEC (September 2013), *GEC Learning guidance on outcome targets*.

What impact has the SPW had on literacy?

Table 18: Difference in literacy scores between baseline and endline in treatment and control groups

Literacy		Avanti[1] Group A (Full intervention)	Group B (No ICT labs)	Group C (No stipends)	Nairobi (Kenya)	Ghana	Nigeria	Coca-Cola[5] Cycle 1[3]	Cycle 2[4]
		Kenya			Nairobi (Kenya)	Ghana	Nigeria	Nigeria	
Test/ language		Uwezo (English)			Uwezo (English)			EGRA (English)	
Unit		Standardised score for wpm			/40			wpm	
Control group used		Yes			Yes			Yes	
In-school cohort	Baseline	Grade 3			Grade 4			Grades 7-11	
	Endline	Grade 6			Grade 6			N/A[6]	
Out-of-school cohort		No			No			No[2]	
Overall impact (all grades)									
Intervention	Baseline	-0.02	0.06	0.04	25.3	14.2	4.9	64.0	
	Endline	0.09	-0.01	-0.07	32.4	22.2	10.7	80.1	
Control	Baseline	-0.09	-0.09	-0.09	26.2	12.5	3.6	64.9	
	Endline	-0.02	-0.02	-0.02	33.2	20.6	7.6	76.1	
DID		0.04	-0.13	-0.18	-0.34	-1.6	1.6	4.75	
P value		0.830	0.570	0.440	0.569	0.137	**0.077**		
Target (SD)		0.4	0.4	0.4	0.2	0.2	0.2	0.2	
Literacy impact (SD)		0.04	-0.13	-0.18	-0.07	-0.12	0.14	0.10	
Achievement against target		-20%			-35%	-61%	71%	52%	
		Negative achievement			Negative achievement	Negative achievement	Half- achieved	Half- achieved	N/A

[1] Results for Avanti are insignificant due to small sample sizes.

[2] The learning outcome is measured only for in-school girls. However the project also works with out-of-school girls who take part in ENGINE clubs and count as beneficiaries (hence the numbers sum up ISG and OSG).

[3] The datasets are not merged. Cohort was not tracked in Cycle 1, so the datasets are not merged. No regression possible as evaluator did not merge baseline and endline data (different evaluator).

[4] For Cycle 2, Coca-Cola literacy data is inconclusive due to non-recoverable gaps and inconsistencies in the data.

[5] For Cycle 1, Coca-Cola tested proportionally more girls from grade 11 at endline than at baseline (67% against 40%), especially from treatment group. Therefore, a weighted average for baseline and endline scores based on the sample composition at these two data collection points would artificially inflate endline value for literacy and numeracy scores and conduce to higher difference-in-difference scores. Instead, for both Cycles, and as the FM also did, the weighted average we computed here for baseline and endline for treatment and control groups are solely based on the composition of endline sample.

[6] No information about the cohort girls' grades at endline has been provided in the documents the EM accessed. Endline evaluation occurred 9 months after the baseline one so we can assume that these girls are still in grades 7-11 at endline if not in grades 8-12.

Key: Grey: Inconclusive N/A

Notes:

The first rows of the table describe the **type of assessment** used in order to measure literacy. The second row refers to the **unit** used for the test, the third, the presence or absence of a **control group** in the project analysis. The fourth row indicates the **grades** of the girls tested, and the fifth row details the presence or absence of an **out-of-school cohort** tested.

The following rows show the **baseline and endline scores for the treatment and control groups** extracted from the projects' outcome spreadsheets. For Coca-Cola, we computed weighted averages of the different grades tested. The **adjusted difference-in-difference (DID)** row is an estimate of the impact of the project on treatment group compared to the control group. It has been calculated by the FM based on adjustments that are specific to each project and the **p-value** row just below indicates if the impact is statistically significant (we consider that it is if the p-value<0.100 highlighted in bold in the table when it is the case).

Endline target in standard deviation (SD) shows the project's midline target in standard deviation, which is the same for both literacy and numeracy. The **literacy impact** row shows the impact achieved by each project in standard deviations. **Project achievement against target** is the percentage of target achieved when we compare **literacy impact in SD** to the target in SD. The last row is a translation in qualitative terms of the adjustment impact in percentage. It is rated from 0 to 5 (FM rating): 0 – Inconclusive or N/A; 1 – Negative achievement (<0%); 2 – Target not achieved (0-50%); 3 – Target half-achieved (50-100%); 4 – Target achieved (>100).

In the section below on project literacy scores, we used the raw data provided by the project to analyse and represent literacy and numeracy average scores at baseline and endline, as well as their distributions. This explains why some small – but negligible – discrepancies can be found between the average scores presented in the graphs below, and the average scores presented in Table 18 which were derived from the projects' outcome spreadsheets.

Avanti

To identify patterns of change in literacy for Avanti, we chose not to analyse average standardised scores as presented in Table 18, but words per minute (wpm) scores. Although impact measured from the wpm scores can differ from the impact observed in standardised scores, this choice allows us to observe girls' improvements by using a concrete measure[49], and to compare their literacy skills to international benchmarks.

From baseline to endline, literacy levels improved in all treatment groups and in the control group. Furthermore, average wpm scores show that girls reached by Avanti were not particularly marginalised in terms of literacy levels.

At baseline, the average words per minute in the three treatment groups was 46 wpm across the whole sample, which is one word per minute above the basic comprehension level standing at 45 wpm according to international benchmarks[50]. It is also worth noting that the girls that were in Grade 3 at baseline had an oral reading fluency of a Grade 2 child (45 wpm). At endline, we found that girls in treatment Groups A (full intervention) and B (not ICT labs), as well as in the control group, who passed through to Grade 6, had an oral reading fluency corresponding to a Grade 4 child (85 wpm). Children in treatment Group C (no stipends) were not far behind with 81 wpm on average. This means that **after two years (see** Section 2.3.2, Table 9**), these girls improved their literacy level by two grades**. Despite the consistency with international benchmarks, it is important to note that, in the context of the GEC, these girls do not appear as being particularly marginalised in terms of their literacy levels.

Despite the improvement observed from baseline to endline and the positive but negligible effect in Group A (full intervention), the Avanti overall had no effect on reading fluency.

Despite the improvement observed during the project's implementation, girls' literacy levels in the treatment groups improved less than those in the control groups. This negative effect was also found when we observed the impact in standard deviation that derives from standardised wpm scores (0.08 SD), but it is not statistically significant. We observed the same pattern when we compared each of the treatment Groups A, B and C to the control group separately - although for Group A, the impact in standard deviation terms is negligible (0.04 SD) as Table 18 shows. None of these results have been found to be significant at 10%.

Figure 16: Avanti - Average literacy levels (in wpm) between baseline and endline, treatment Groups A, B and compared to the control group
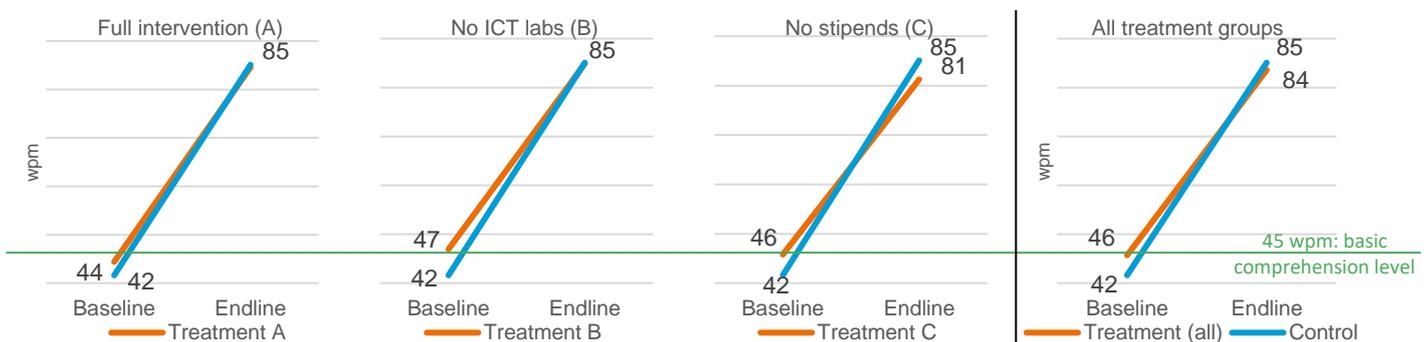
Stipends seem to have a positive effect on girls whose initial literacy skills were very weak, whereas ICT labs seem to affect girls that already had some literacy skills to reach higher levels. However, these effects were not statistically significant.

---

[49] By concrete, we mean that we can link the score to an aptitude. For instance, in the standard Uwezo test, reaching level 2 in numeracy means that the child can count numerical numbers from 1 to 9. Similarly, for wpm this refers to the ability to read with more or less difficulty. Alternatively, standardised scores do not provide us such information on the child's performance.
[50] 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.

As shown in Annex D.1.1, at baseline, many girls were only able to read less than 10 wpm (21%, 10%, 9% and 15% respectively in Groups A, B, C and D). At endline, about 80% of the girls from each group were able to read at least 60 wpm, and between 10% and 15% of the girls could read 108 wpm or more, which corresponds to the minimum average level for Grade 6 according to international benchmarks.

In the groups where ICT labs or stipends were not provided with the rest of the intervention (Groups B or C), the project had no effect. In treatment Groups A and B **where stipends were provided, the proportion of girls having very weak reading skills (less than 40 wpm), decreased more** than when stipends were not provided (Group C). We can therefore assume that stipends possibly allow the most marginalised girls (those who are more in need and had lower attendance in school) to improve their oral reading fluency. On the other hand, **when ICT labs are provided (Groups A and C), more girls tended to reach higher scores** (more than 90 wpm) than in a situation where there are no such labs. Therefore, we can assume that ICT labs would possibly allow girls that are not as marginalised to improve their oral reading fluency. No statistical test has been run, and therefore we cannot confirm that these effects are attributable to the project.
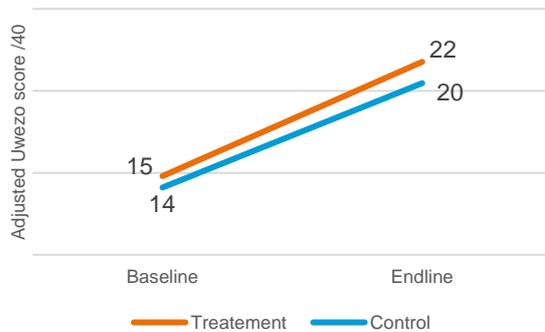
Discovery

Unlike Avanti, Discovery covers more than one geographic area – four areas in three countries with very different social and cultural contexts. In Nairobi (Kenya), Ghana and Nigeria, Discovery conducted a quasi-experimental evaluation) by sampling treatment and control groups for each. For Wajir (Kenya), where Discovery did not find enough eligible schools to conduct such an evaluation, a performance evaluation was carried out instead. Nevertheless we can still present literacy scores in Wajir (Kenya) for Round 1 and Round 2 in order to assess whether the length of the project's implementation had an effect on literacy outcomes.

Across Discovery's target areas, literacy scores improved slightly more in treatment than in control groups.

Overall, in all locations where Discovery conducted a quasi-experimental evaluation – i.e. in Nairobi (Kenya), Ghana and Nigeria – a negligible positive effect of 1 word per minute was observed from baseline to endline. No statistical test has been run in order to verify the significance of this effect (i.e. the attribution to the project).

Figure 17: Discovery – Average literacy levels across the areas covered by the project, between baseline and endline, treatment compared to the control group


Compared to control groups, the project had no impact on the average literacy levels in the treatment groups

On average, Discovery's effect on literacy is close to zero: the difference-in-difference scores for the whole project including for Nairobi (Kenya), Ghana and Nigeria equals 0.51 points on the adapted Uwezo scale. When data was disaggregated by country, we observed negative results in Nairobi (Kenya) and Ghana (-0.07 SD and -0.12 SD respectively), although these were not statistically significant. By contrast, in Nigeria, the project had positive effects which were statistically significant at 10% (impact in SD=0.14, p-value=0.077). This might suggest that the initial situation of the girls (who achieved lower literacy scores) allowed the project to have more of a positive impact and demonstrate improvement from baseline.

In Wajir (Kenya), the length of implementation was not associated with a change in the reading scores, which decreased by one point from Round 1 to Round 2 for Year 1 and Year 2 schools. Also, the changes observed were equal to -0.09 SD and 0 SD respectively and so the project did not meet the target of 0.2 SD.

Figure 18: Discovery - Average literacy levels between baseline and endline, treatment vs. control groups
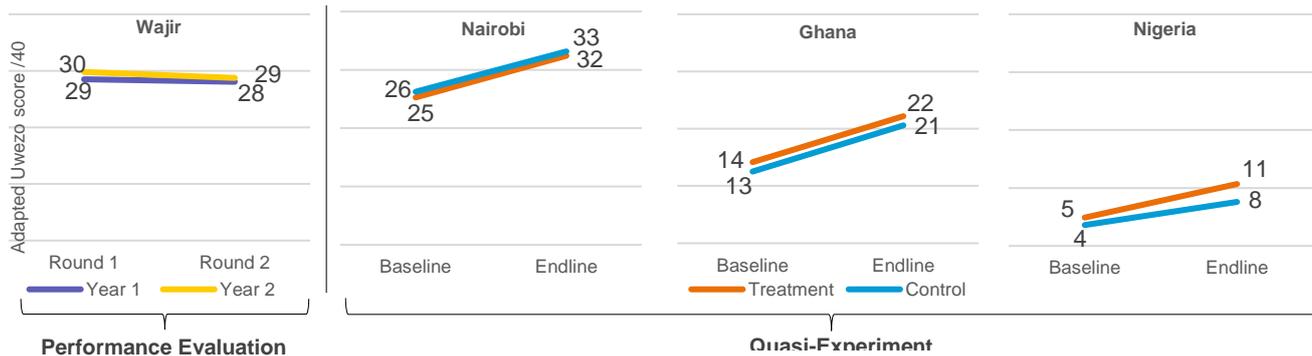
Discovery had a positive and statistically significant impact on literacy skills only in the areas where girls' performance was initially (i.e. at baseline) low.

In Nigeria, the programme appears to have reduced the proportion of girls who had scores close to zero at baseline. In Ghana, where intermediate scores were observed, the effect (statistically insignificant) is null, whereas in Kenya, where scores were around 25-30, the effect is slightly negative (but still statistically insignificant).

There are a number of possible explanations for the finding that the programme only had a positive effect in Nigeria. On the one hand, since statistical tests were conducted at the 10% level of significance, it is possible the result is a false positive. On the other hand, if this trend is an accurate reflection of the programme's effectiveness, it could suggest that Discovery's interventions were best suited to students starting with very weak literacy skills. But since the programme did not focus on remedial education or basic reading skills it isn't obvious why this would be the case. Another explanation could be that the programme was more effective in filling gaps in a weak education system than in supporting school systems that were already moderately effective. The much weaker literacy scores in Nigeria could reflect problems with the school system that were not as important in Kenya or Ghana. Other differences in the context and implementation across Discovery's project areas are explored in Section 3.4.

Several other trends can be observed:

- **In Wajir (Kenya), the length of the Discovery programme had no impact on girls' literacy skills.** Between baseline and endline, we observed a decline both in the proportion of low skilled and high skilled girls in Year 1 and Year 2 groups (see Annex D.2.1.1), the overall effect leads to a slight decline in the average score. Because there is no control group for this project area, it is not possible to draw definite conclusions about whether the project is responsible for the effects on either the high or low skill groups. However, this trend is consistent with the evidence from the other three projects that most positive effects were concentrated among the lowest performing students. The length of implementation did not seem to play any role in causing the effects observed that were similar in both Year 1 and Year 2 groups.

- **In Nairobi (Kenya), improvements were observed in literacy skills, but the project did not demonstrate any impact.** Although average baseline scores were lower than in Wajir (Kenya), most girls from both groups improved and reached higher scores at endline: 62% of girls from the treatment group and 64% of girls from the control group scored 32 points out of 40 or more (see Annex D.2.1.2).

- **In Ghana, the general improvement observed in literacy - with numerous high scores reported at endline – is not attributed to Discovery's project intervention.** As shown in Annex D.2.1.3, at baseline, half of the girls tested in both treatment and control groups could not read (i.e. they scored less than 8 / 40 points on the adapted Uwezo scale). However, 12% of girls in both groups also reached high scores (they scored 32 points or more). At endline, we still observe these two distinct subgroups but there were fewer unskilled girls (19% and 26% from treatment and control groups scoring less than 8 points) and more skilled girls (31% of girls from both groups scoring 32 points or more). Overall, girls from treatment groups did not improve more than girls from control groups.

- **In Nigeria, where most of the girls surveyed had scores close to zero at baseline, a general improvement has been observed in both treatment and control groups.** These improvements were higher in treatment groups. At baseline, in Nigeria, more than 70% of girls in both groups of the girls surveyed were unable to read (they scored less than 4 points out of 40 on the adapted Uwezo scale). At endline, most of these girls were still illiterate but their proportion plunged to 45% and 54% of girls from

treatment and control groups respectively. Improvements were more evident in the treatment group where the project demonstrated a positive and significant impact, though in absolute terms this was still small, only 2 words per minute (see Annex D.2.1.4).

Coca-Cola

Although Coca-Cola targeted in-school and out-of-school girls, only the former have been tested on their learning outcomes. Indeed, out-of-school girls were mainly trained in business skills, entrepreneurship, leadership skills, and financial education, but not specifically in literacy and numeracy skills. Therefore, testing them was not perceived as relevant or appropriate.

In Cycle 1, it is worth noting that there were some issues with the administration of the tests, however, the FM advised that this should not interfere with the overall final analysis[51].

For Cycle 2 baseline and endline, literacy data was found to be unusable, therefore we did not present any analysis in this report.

For Cycle 1, literacy data was available at baseline and endline. However, no statistical test was carried out to measure the project's attribution. As previously explained, Coca-Cola appointed a new external evaluator after the Cycle 1 baseline research was carried out. The new external evaluator disregarded the baseline results and data of the previous external evaluator (as noted, this is because these may have been exaggerated and seemed unrealistic)[52].

Unlike the other projects that chose to test only one grade, Coca-Cola tested girls from Grade 7 to Grade 11. However in Cycle 1, samples for Grades 7 and 8 were too small to provide reliable information on girls' average literacy skills. Although these results are not presented in this report, they are accounted for in the total score in Figure 19.

From baseline to endline, oral reading fluency improved for Cycle 1 in both treatment and control groups, as well as for all of the grades tested. On the whole, the Cycle 1 increase has been greater in the treatment group than in control group.

At baseline, in Grades 9 to 11, oral reading fluency was above the comprehension level (45 wpm)[53]. At endline, oral reading fluency increased to 74 wpm, 75 wpm and 82 wpm in treatment groups for Grades 9, 10 and 11 respectively.

Compared to international benchmarks (see Figure 14), 16-19 year old in-school girls tested by Coca-Cola have the oral reading fluency age of 7 to 9 year old children in Grades 2-3 (45 wpm and 63 wpm respectively). At endline (i.e. 9 months after the intervention (education cycle)), cohort girls barely improved by one grade: most of the girls had Grade 3 level oral reading fluency (63 wpm) and none of the sub-groups reached Grade 4 level (85 wpm).

Improvement in treatment groups compared to control groups was observed as the grades got higher: negative in Grade 9 (- 4 wpm) and Grade 10 (-1 wpm), becoming positive in Grade 11 (5 wpm). The overall improvement observed in treatment over the control group stood at 7 wpm and was mainly drawn from Grade 11 girls, which also represent two thirds of the total sample at endline. However, as explained in Box D.2, no statistical test/ regression has been run in order to verify the significance of these results (or the attribution of the project's intervention).

---

[51] Administration errors have been observed at endline but, as they are equally likely in both treatment and control groups, and as the generated spike is observed in both treatment and control group distribution shapes, the FM approved the use of these data to make comparisons in the oral reading fluency between baseline and endline, and between treatment and control groups. Nonetheless, it has to be noted that in the absence of these errors, the distributions would probably be more skewed to the right with higher levels of wpm for about one quarter of the girls tested.

[52] The external evaluator justified this by providing examples, such as that in the Federal Capital Territory, the average wpm measured at baseline was above 90 wpm and during endline it was significantly reduced to 45 wpm on average.

[53] 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.

Figure 19: Coca-Cola, Cycle 1 – Average literacy levels (wpm) for grades 9 to 11 at baseline and endline, treatment vs. control group
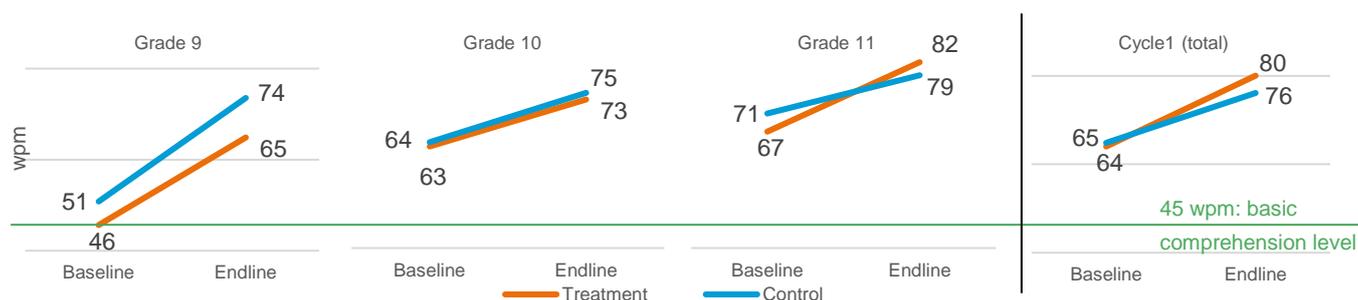
Observation of wpm distributions show that during Cycle 1, Coca-Cola has probably not contributed towards decreased illiteracy.

At baseline we observed that many girls had low scores (see Annex D.3.1): one third of the girls from treatment groups and control groups scored less than 10 wpm. At endline, there was a drop in illiteracy especially in the treatment group (8% of the treatment group and 10% of the control group).

3.2.1 What impact has the SPW had on numeracy?

Table 19: Difference in numeracy scores between baseline and endline in treatment and control groups

Numeracy			SPW projects by country and region					
		Avanti[1]			Discovery			Coca-Cola[2,5]	
		Group A (Full intervention)	Group B (No ICT labs)	Group C (no stipends)				Cycle 1[3]	Cycle 2[4]
		Kenya			Ghana		Nigeria		
Test		Uwezo			Uwezo			EGMA	EGMA
Unit		Standardized Uwezo scores			/40			/100	/100
Control group used		Yes			Yes			Yes	Yes
In-school cohort	Baseline	Grade 3			Grade 4			Grades 7-11	
	Endline	Grade 6			Grade 6			N/A[6]	
Out-of-school cohort		No			No			No[2]	
Overall impact (all grades and out-of-school girls)									
Intervention	Baseline	0.10	0.10	0.10	26.3	22	4.8	56.1	50.0
	Endline	0.39	0.27	0.14	33	21.6	7.9	57.5	59.4
Control	Baseline	0.10	0.10	0.10	25.8	20.9	4.5	52.9	54.2
	Endline	0.23	0.23	0.23	33.4	21.5	6.2	58.1	60.5
DID		0.16	0.04	-0.09	-1.35	-1.80	0.80	-3.67	1.60
P value		0.290	0.820	0.630	**0.069**	**0.068**	0.305		0.470
Target (SD)		0.4			0.2			0.2	
Impact (SD)		0.16	0.04	-0.09	-0.19	-0.17	0.08	-0.19	0.07
Achievement against target		8%			-96%	-83%	41%	-95%	34%
		Not achieved			Negative achievement	Negative achievement	Not achieved	achievement	Not achieved

[1] Results for Avanti are insignificant due to small sample sizes.

[2] The learning outcome is measured only for in-school girls. However the project also works with out-of-school girls who take part in ENGINE clubs and count as beneficiaries (hence the numbers sum up ISG and OSG).

[3] The datasets are not merged. Cohort was not tracked in Cycle 1, so the datasets are not merged. No regression possible as evaluator did not merge baseline and endline data (different evaluator).

[4] Partially conclusive due to high attrition.

[5] For Cycle 1, Coca-Cola tested proportionally more girls from grade 11 at endline than at baseline (67% against 40% at baseline for Cycle 1), especially from treatment group. Therefore, a weighted average for baseline and endline scores based on the sample composition at

these two data collection points would artificially inflate endline value for literacy and numeracy scores and conduce to higher difference-in-difference scores. Instead, for both Cycles, and as the FM also did, the weighted average we computed here for baseline and endline for treatment and control groups are solely based on the composition of endline sample.

[6] No information about the cohort girls' grades at endline has been provided in the documents the EM accessed. Endline evaluation occurred 9 months after the baseline one so we can assume that these girls are still in grades 7-11 at endline if not in grades 8-12.

Notes

The first rows of the table describe the **type of assessment** used in order to measure literacy. The second row refers to the **unit** used for the test, the third, the presence or absence of a **control group** in the project analysis. The fourth row indicates the **grades** of the girls tested, and the row five details the presence or absence of an **out-of-school cohort** tested.

The following rows show the **baseline and endline scores for the treatment and control groups** extracted from the projects' outcome spreadsheets. For Coca-Cola, we computed weighted averages of the different grades tested. The **difference-in-difference (DID)** row is an estimate of the impact of the project on treatment group compared to the control group. It has been calculated by the FM based on adjustments that are specific to each project. This has been retrieved from the outcome spreadsheets and the **p-value** row just below indicates if the impact is statistically significant (we consider that it is if the p-value<0.100 highlighted in bold in the table when it is the case). **Endline target** in standard deviation (sd) shows the project's midline target in standard deviation, which is the same for both literacy and numeracy. The **literacy impact** row shows the impact achieved by each project in standard deviations. **Project achievement against target** is the percentage of target achieved. The last row is a translation in qualitative terms of the adjustment impact in percentage. It is rated from 0 to 5: 0 – Inconclusive or N/A; 1 – Negative achievement (<0%); 2 – Target not achieved (0-50%); 3 – Target half-achieved (50-100%); 4 – Target achieved (>100%).

Avanti

Presentation of absolute scores does not allow comparison between baseline and endline in order to measure a project's impact. Therefore, we use standardised scores in this section.

Avanti's intervention had no effect on numeracy on treatment Groups A (full intervention) and B (no ICT labs)..

Overall, Avanti's intervention had a positive but negligible effect on treatment groups over the control groups (with an impact equal to 0.03 SD). The result is not statistically significant. In treatment Group A (full interventions) we observed the highest improvement (0.16 SD). It is followed by treatment Group B where stipends were offered but without ICT labs (0.04 SD). Lastly, we observed a negative effect on numeracy for treatment Group C, where ICT labs were provided but not stipends (with an impact equal to -0.09 SD). However, none of these results were statistically significant, and therefore cannot be attributed to the project.

Figure 20: Avanti: average numeracy levels (in % of correct responses) between baseline and endline, Groups A, B and C compared to the control Group D

The provision of ICT labs seems to have allowed the least skilled girls in numeracy to improve their skills, whereas stipends seem to have allowed the girls who were initially better at numeracy skills to improve.

Distributions of the scores shown in Annex D.1.2, demonstrate that the full intervention (Group A) allowed the **girls that reached the lowest scores at baseline to improve** at endline compared to the control group. This effect was also observed in the presence of **ICT labs** (Group C). Full intervention also **helped more girls to reach higher scores** at endline, which is also the case in the group where **stipends** have been provided (Group B). These results are not statistically significant and therefore cannot be strictly attributable to the interventions. The project report also shows that students who had used Maths-Whizz for at least the recommended amount of time (30 minutes of weekly usage) advanced noticeably more than the group as a whole. However, these results should be interpreted with care due to a small sample of students meeting the recommended usage (11 students from the whole sample) and the number of students having a weekly usage of 15 min or more (40 students).

Discovery

Across Discovery's implementation areas, numeracy scores improved for both treatment and control groups similarly.

With no statistically significant effect on numeracy, the improvements observed from baseline to endline cannot be attributed to Discovery's project.

Figure 21: Discovery – Average numeracy levels across the areas covered by the project, between baseline and endline, treatment compared to the control group
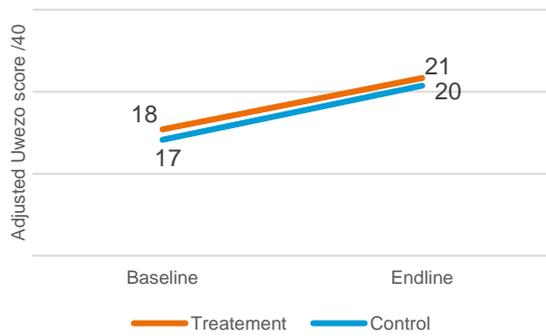

Between baseline and endline the average numeracy scores have slightly improved or stagnated.

As observed for literacy skills, numeracy scores vary considerably from one country to another, but in all cases little improvement was observed from baseline to endline. Average scores stood at about 30 points out of 40 in Wajir and Nairobi (Kenya), 20 points in Ghana, and less than 10 points in Nigeria at baseline and endline. With the exception of Nairobi (Kenya), the effects observed are very small.

Looking at the difference-in-difference scores, in Wajir (Kenya), we identify a small, positive association between the length of implementation and the numeracy score (difference-in-difference score stood at 0.42 in Year 1 and 0.41 in Year 2). In Nairobi (Kenya) and in Ghana, Discovery had a negative effect (with impacts expressed in standard deviations equal to -0.19 SD and –0.17 SD respectively) that is statistically significant. In absolute terms this was a very small difference. By contrast, in Nigeria, similar to literacy, the project had a positive albeit not statistically significant effect on numeracy (0.08 SD).

Figure 22: Discovery: average numeracy levels between baseline and endline, treatment compared to the control group (four locations)
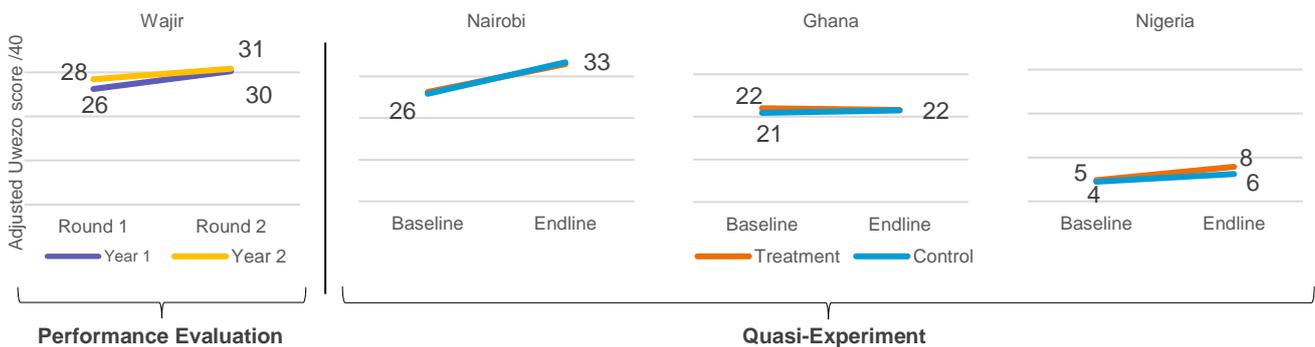
By using in-depth analysis to examine the distribution of effects, it appears that Discovery had a negative and statistically significant effect where girls had initially good numeracy skills, and had no effect where girls were initially unskilled.

This was observed in all Discovery's locations with the exception of Wajir (Kenya) where numeracy skills of girls were observed to improve the longer the implementation of Discovery's interventions continued.

As suggested in the literacy findings, **Discovery appears to have had a negative effect in contexts where the education system is functioning relatively well** resulting in a deterioration in girls' numeracy skills compared to

the control groups. Nevertheless, except in Nairobi (Kenya), all the effects on numeracy observed in Discovery's projects are very small.

- **In Wajir (Kenya), the duration of Discovery's project implementation had a positive effect on numeracy skills.** In Round 1, few girls scored low in numeracy. In Round 2, the skills of these girls did not seem to improve. On the other hand, girls that scored 20 points or more out of 40 improved their numeracy skills. This effect was more evident in the Year 1 group who were exposed to programme activities longer than those in the Year 2 group, which suggests that the length of Discovery's programme implementation had a positive effect on numeracy skills (see Annex D.2.2.1).

- **In Nairobi (Kenya) extensive improvements in numeracy scores cannot be attributed to the programme.** From baseline to endline, the distribution of numeracy scores presented in Annex D.2.2.2 shows an overall improvement for all of the girls achieving high scores at endline (70% of girls from both groups reached a score that is higher or equal to 32 points out of 40 on the adapted Uwezo scale). However, girls in the control groups improved more than the girls in the treatment groups, and this difference is statistically significant.

- **In Ghana moderate improvements are observed in numeracy, but more so for control than treatment.** At baseline, Ghana's sample was composed of girls achieving different levels of numeracy (see Annex D.2.2.3). At endline, we observed an improvement of the girls who were least skilled at baseline, but also a decrease in the performance of the most skilled girls, especially in the treatment group. This suggests that the Discovery programme had a negative effect on the girls that had better numeracy levels at baseline.

- **In Nigeria the proportion of low skilled girls declined from baseline to endline, and this improvement was slightly higher in treatment than in control groups.** At baseline, most of the sample scored less than 4 points out of 40 on the adapted Uwezo scale. At endline, there are still a lot of girls who are unskilled in numeracy, but the proportion of girls scoring less than 4 points dropped from 62% to 52% in treatment groups, and from 52% to 42% in control groups. Overall, improvements have been slightly better in treatment than in control groups, but not enough to conclude that Discovery had a statistically significant positive effect on numeracy outcomes in Nigeria.

Coca-Cola

The effect of Coca-Cola's programme on numeracy has been negative in Cycle 1, and close to zero in Cycle 2, but the effects are not statistically significant.

In Cycle 1, from baseline to endline, girls' numeracy scores improved less for treatment groups than control groups (with an impact that stood at -0.19 SD), but no statistical test was conducted and therefore attribution to the project is not possible. Similarly, in Cycle 2, the impact was found to be positive (0.07 SD), but it is not statistically significant[54].

Figure 23: Coca-Cola: average numeracy scores between baseline and endline, treatment compared to the control group in Cycle 1 and Cycle 2


---
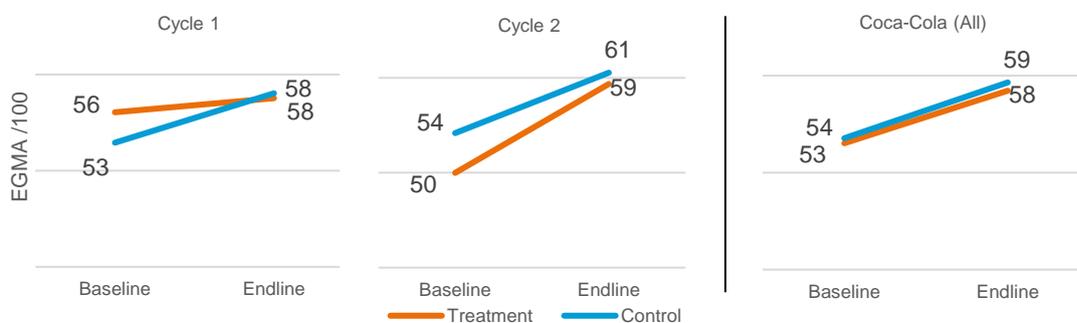
[54] From Figure 23, for Cycle 2, we would have expected a higher impact. However, as attrition was high, the DID regression led to finding a lower adjusted difference-in-difference equal to 1.6.

Coca-Cola's Cycle 1

In Cycle 1, Coca-Cola's effect decreased as the grades of the girls increased.

In Cycle 1, the projects' positive effect decreased as the grade years of the girls increased. In Grades 9 and 10, there was an improvement in treatment groups over control groups that stood at 2 and 3 EGMA points respectively. By contrast, in Grade 11 the treatment group average score declined and the effect is negative (-5 EGMA points).

Figure 24: Coca-Cola, Cycle 1: average numeracy scores between baseline and endline, treatment compared to the control group
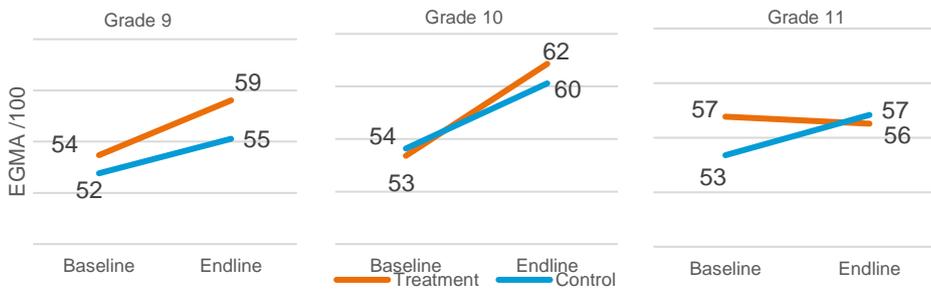
In Cycle 1, Coca-Cola had an overall negative effect on numeracy with a decrease in numeracy scores among the least skilled girls, but the effect was not statistically significant.

Despite a general improvement observed in the distribution of EGMA scores as presented in Annex D.3.2.1, there was also an increase in the proportion of low scores observed, especially in treatment groups. This could suggest that Coca-Cola had a negative effect on numeracy, but the effect was not statistically significant.

Coca-Cola's Cycle 2

In Cycle 2, Coca-Cola's effect is close to zero.

In Cycle 2, the effect is positive but not statistically significant. Scores observed by grade do not permit us to identify a clear pattern for the project's effect: in Grades 9 and 11 girls in the treatment group improved less than control groups whereas in Grades 8 and 10 the treatment group improvement over the control group is positive. However, these differences are not statistically significant, and taken together this evidence is consistent with the hypothesis that Cycle 2 had no effect on numeracy.

Figure 25: Coca-Cola, Cycle 2: average numeracy scores between baseline and endline, treatment compared to the control group
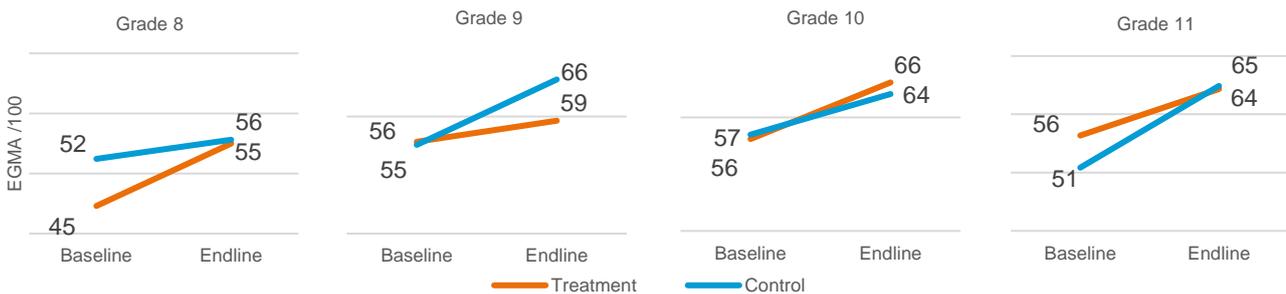
In Cycle 2, Coca-Cola had no effect on numeracy.

In Cycle 2, we observe improvements in the scores distributions (see Annex D.3.2.2) from baseline to endline. Although the proportion of girls reaching higher scores is smaller in treatment group than in control group at baseline and endline, the gap between both groups is only slightly reduced during the project implementation which suggests an **absence of impact of Coca-Cola's project on numeracy**.

<div style="border: 1px solid #29ABE2; padding: 10px;">

Lessons learned

- **Projects had difficulties in improving learning outcomes for girls during the lifetime of the programme**. Targets were set at baseline in a way that accounted for the specific learning distributions among target populations, and hence reflected the context in which each project intervenes. A standard deviation of 0.2-0.4 is also rather low as a target compared to other educational programmes, which suggests that projects should have been able to reach their targets. However, it is possible that since projects aimed to target *marginalised* girls, learning improvements for this specific population are more difficult to secure and take more than two years of support to materialise.

- **The lower the initial learning outcomes, the higher the chances to observe a positive significant impact**. Indeed, in a context in which the educational system is completely deficient, any additional intervention can potentially contribute to improve educational outcomes. Furthermore, where learning levels are very low, even small differences between treatment and control groups can be statistically significant. This situation has been observed in Discovery's project for literacy where the only positive impact observed was in Nigeria. By contrast, in areas where the education system appears to be performing relatively well, an intervention competing with the existing system can have adverse effects as was the case for numeracy in Ghana and Kenya for Discovery's project.

- **The time spent on online educational platforms seems to be related to higher improvements in numeracy**. However, Avanti's project did not ensure that a sufficient number of beneficiaries would spend at least the recommended amount of screen time – 30 minutes – to improve educational outcomes. This issue is related to the small number of computers provided by Avanti to each school – 25 for 1,000 pupils on average.

- **Two hours of tutoring per week for nine months during which literacy and numeracy is taught for a very small amount of time was not sufficient to ensure improvements in the GEC's educational outcomes**. In both cycles, Coca-Cola did not enable in-school beneficiaries to progress their literacy and numeracy. By contrast, for out-of-school girls, the time made available was sufficient to observe improvements in financial literacy.

- **Providing a combination of stipends and ICT labs has more effect on literacy and numeracy than just providing stipends or ICT labs alone**. The group that only received stipends improved the least in terms of their literacy and numeracy. However, it should be noted that this group was composed of girls that were specifically targeted on the basis of educational marginalisation, specifically their lack of capacity to afford the cost of schooling. Therefore comparisons between these three groups should be treated with care.

</div>

3.3 What impact has the SPW had on enabling marginalised girls to be in school?

<div style="background: #29ABE2; color: white; padding: 10px;">

Key Findings

Across the SPW, there were issues with the attendance data and measurement of attendance. These limited our ability to generate conclusive findings.

At endline, **Avanti's** attendance results were inconclusive. School attendance records for control schools were not available or not properly kept. For **Coca-Cola**, the attendance data provided was collected at the Safe Spaces rather than in schools – they were calculated based on an average of attendance rates during the nine months of the course cycle. **Discovery** achieved its attendance target at endline in Nairobi (Kenya) and Nigeria, whereas **Coca-Cola** nearly achieved it in their Safe Spaces for both its Cycles.

</div>

To assess the extent to which girls are in school across the SPW projects, we looked at **attendance** data that projects recorded in their Outcome Spreadsheets. The attendance data that projects reported was calculated based on an average of the school days attended. Projects collected data using: a digital attendance monitoring system (Avanti); school registers (Discovery); and Learning Space Centre registers (Coca-Cola). For Avanti and Discovery, data was cross-checked through spot checks, as well as through other qualitative (Focus Group Discussions and Key Informant Interviews) and quantitative (household surveys) research methods.

Attendance rates measured by SPW projects are compared to a target set in the project logframes that correspond to a defined rate (Coca-Cola) or a percentage increase in the treatment group over the control group (Discovery). Attendance rates are also compared to the national female net attendance ratio in primary schools provided by the Demographic Health Survey (DHS) for Kenya (2014), Ghana (2014) and Nigeria (2013).

> *Enrolment across the SPW*
>
> *The GEC requirements did not specify that SPW projects should provide enrolment information unless the project had a specific focus on interventions aimed at improving enrolment and/or retention. In their Outcome Spreadsheets neither Avanti nor Coca-Cola provided data on enrolment, whereas Discovery measured enrolment rates that were based on their household surveys, but these results were not accurate because the households were not randomly selected. Therefore, no data for enrolment is presented in this report. However, for Coca-Cola, we have information on educational cycles' completion rates for out-of-school girls that we briefly present in this section.*

Table 20: Average attendance rate between baseline and endline, treatment compared to the control across the SPW projects[55]

Attendance		SPW projects by country and region							
		Avanti			Discovery			Coca-Cola[1]	
		Group A	Group B	Group C				Cycle 1	Cycle 2
		Kenya			Kenya	Ghana	Nigeria	Nigeria	
Source		Attendance monitoring System			School registers			Safe Space Registers	
Intervention	Baseline	80%	79%	77%	88%	92%	88%		
	Endline	81%	79%	74%	97%	93%	96%	56%	63%
Control	Baseline				94%	93%	94%		
	Endline				97%	94%	97%		
Target					2%	2%	2%	75%	75%
Achievement against target					255%	5%	110%	74%	84%
		N/A			Achieved	Not achieved	Achieved	Half-achieved	Half-achieved

[1] For Coca-Cola, reported attendance rate at the Safe Spaces concerns in-school girls. However, the project also provided data for out-of-school girls that will be analysed further in the section.
Key: Grey: Inconclusive or not available

Avanti

> **Box 4: Attendance data collection and target for Avanti**
>
> Attendance for treatment Groups A, B and C schools (195 schools) was recorded on a daily basis through an electronic attendance monitoring system (sQuid) introduced by the project. This attendance was cross-checked with the schools' registers and through termly spot checks. Attendance data for control schools was collected on a termly basis through the attendance records of the schools, and verified through termly spot checks. However, the project concluded that data for control schools was not reliable because many schools were keeping poor attendance records. Therefore, no data is provided for control schools (Group D).
>
> In treatment schools, baseline data was collected between April and June 2015 when girls were in Grade 4. Endline data was collected between September and November 2016 when the cohort was in Grade 5.
>
> According to the FM, Avanti never targeted a precise improvement in attendance "over and above the control" since they knew that data from the control group would never be fully comparable with the treatment group (which uses sQuid cards). In addition, the control group data was in the end deemed unusable, therefore no target for attendance was set.

---

[55] Only Discovery ran statistical tests to measure the project's impact significance. With no statistical test conducted on attendance data, **we cannot say whether these observed effects are statistically significant**. These findings only reflect the evidence presented by SPW projects (reporting attendance figures).

For Avanti, stipends appear to have prevented a decrease in attendance rates between baseline and endline.

From baseline to endline, attendance rates slightly decreased from 79% to 78% in Avanti's treatment schools.

As shown in Figure 26, at baseline, attendance rates observed in treatment Groups A (full intervention), B (no ICT labs) and C (no stipends) were 80%, 79% and 77% respectively. At endline, it improved by one percentage point in Groups A and B, and decreased by three percentage points in Group C. This might suggest that the stipends improved attendance rates whereas the absence of ICT labs (in group B) did not affect attendance. In all treatment groups, endline attendance rates remained lower than the national average.

As shown in Figure 26 It is worth noting that according to the project endline report, the project observed a gradual decline in 2016 attendance rates and linked it to severe drought observed across the country. Additionally, it noted that the baseline occurred during the second school term, whereas the endline coincided with the third school term which is normally highly disrupted by Kenya Certificate of Primary Education (KCPE) examinations for students in Grade 8, but the examinations affect the whole school between October and November. Although these might be valid explanations for the results, we are unable to independently assess the extent of their effect on the attendance results.

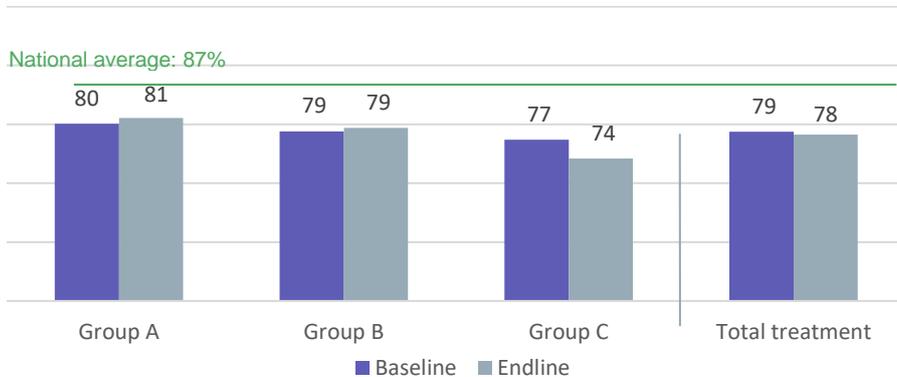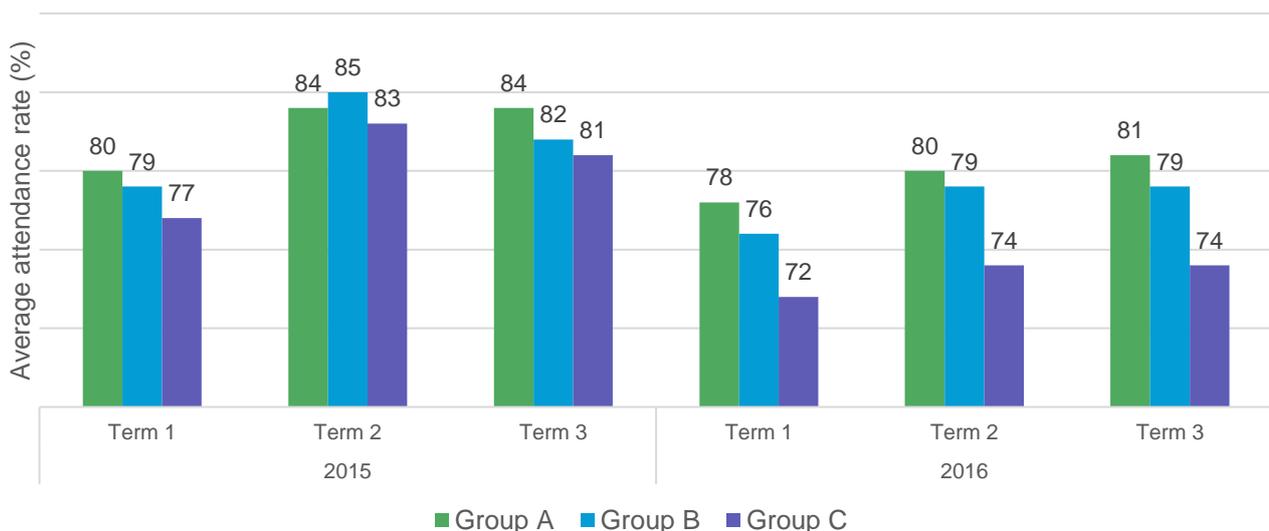**Figure 26: Attendance rates in Avanti intervention Groups A, B and C at baseline and endline**

Figure 27: Avanti average attendance for cohort girls split by intervention between terms

Discovery

> **Box 5: Attendance data collection and target for Discovery**
>
> Attendance data for Discovery was collected in treatment and control schools through school registers, and then triangulated with classroom head counts and household survey data for cohort girls. In all countries, data at baseline and endline was collected during the same school terms. For Kenya and Ghana, the baseline (or Round 1 in the case of Wajir) data was collected between January and April 2014, which corresponded to school term 1 in Kenya, and school term 2 in Ghana. Endline data was collected during the same school terms in 2016. In Nigeria attendance data was collected during school term 3 (May to July) at both baseline (2014) and endline (2016).
>
> For Discovery, the attendance target was set at a 2% improvement from baseline to endline in treatment schools over the control schools. Achievement is calculated through a difference-in-difference estimate. For Wajir, changes in Year 1 schools are compared to changes in Year 2 schools.

For Discovery, attendance rates decreased for both treatment and control areas, although attendance rates were above national averages at both baseline and endline.

On average from baseline to endline, across Discovery project areas – Nairobi (Kenya), Ghana and Nigeria – attendance rates decreased for both treatment and control areas (in treatment groups from 88% to 87%, and in control groups from 91% to 87%). Nevertheless, this average hides a wide range of patterns observed in each of these locations.

As shown in Figure 28 below, in all four areas targeted by Discovery, attendance rates were all above the national averages at both data collection points (i.e. baseline/Round 1 and endline/Round 2).

In **Wajir (Kenya)**, results show that there has been a positive effect on attendance (for both treatment and control groups), the project also achieved its set target, although the result is not significant. From Round 1 to Round 2, results show an increase by 4 percentage points for Year 1 schools, but a decrease by 2 percentage points for Year 2 schools. This represents an overall increase of 5.7% in cohort attendance of Year 1 schools over Year 2 schools, therefore a positive increase in cohort attendance, (albeit weakly associated to the length of the intervention) while exceeding the project's target of 2%. However, the statistical regression does not show a statistically significant effect.

In **Nairobi (Kenya)**, the increase in average attendance from baseline to endline for the cohort girls in the treatment schools (9 percentage points) was three times that of control schools (3 percentage point increase). This result is statistically significant and therefore attributable to the project.

In **Ghana**, the programme's effect is negligible since it increased by about one percentage point in both groups. Therefore, Discovery did not meet its set attendance target.

In **Nigeria**, we observed a decline in treatment school attendance from baseline to endline by 12 percentage points in the treatment group compared to a 14 percentage point decline in the control group. This results in a positive but negligible effect of the project although the target of 2 percentage points has been met. Improvements in the cohort's attendance cannot be attributed to Discovery's project.

Figure 28: Attendance rate in Discovery intervention groups A, B and C at baseline and endline
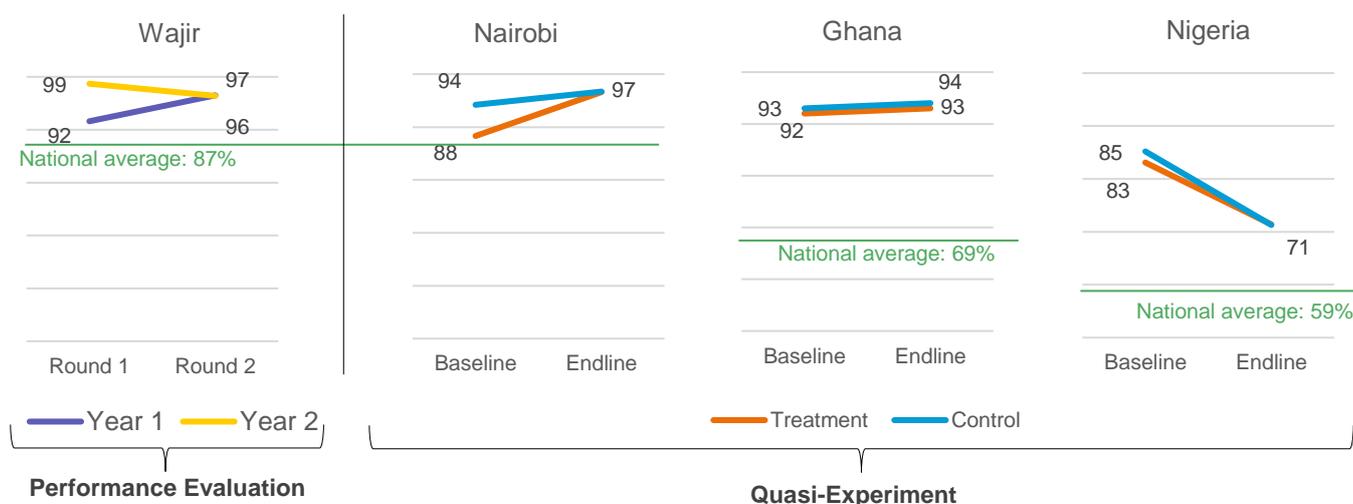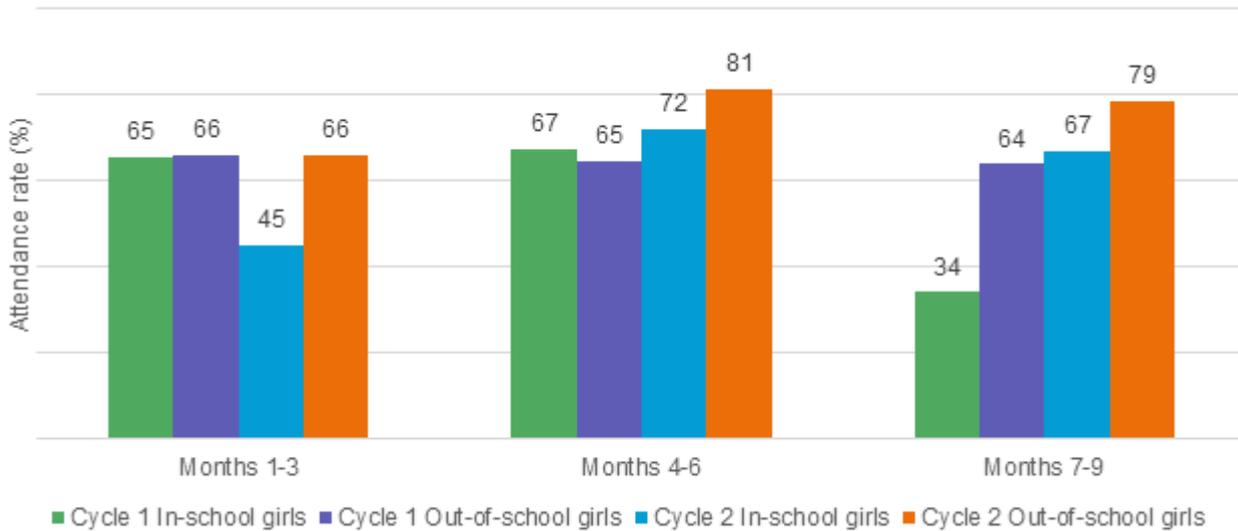

Coca-Cola

> **Box 6: Attendance data collection and target for Coca-Cola**
>
> For the attendance rate outcome, Coca-Cola set a target for the whole duration of the project that corresponds to an average 75% monthly attendance rate. Therefore, only one average monthly attendance rate is provided in Coca-Cola's Outcome Spreadsheets and in Table 20 for each cycle. This target has been achieved at 74% and 84% for Cycles 1 and 2 respectively.
>
> Coca-Cola also collected attendance data for out-of-school girls, and their average monthly attendance rate during the nine month implementation of the project is 65% and 76% respectively in Cycles 1 and 2.

Although Coca-Cola set a target in terms of absolute attendance rate and not in terms of progress in this outcome, and only for in-school girls, we can nevertheless compare the evolution of attendance measured in the Safe Spaces for both Cycles, for in-school girls and out-of-school girls across the nine months of the project's implementation. Data presented in Figure 29 shows various patterns in both cycles, but also between in-school and out-of-school girls, but **overall, attendance has been higher for out-of-school girls, than for in-school girls, especially at the end of the programme implementation** (months 7 to 9). The drop observed for in-school girls' attendance rate at the end of Cycle 1 is absent in Cycle 2. In addition to this, in Cycle 2, we observe an increase in attendance from term 1 (months 1 to 3) and term 2 (months 4 to 6) that was absent in Cycle 1. This suggests that **the project may have rectified the intervention between both cycles**.

Figure 29: Coca-Cola, Cycles 1 and 2 - Average term (three months) attendance rates for in-school and out-of-school girls

■ Cycle 1 In-school girls ■ Cycle 1 Out-of-school girls ■ Cycle 2 In-school girls ■ Cycle 2 Out-of-school girls

!

Enrolment of out-of-school girls in Cycle 1 and 2

The girls targeted by Coca-Cola were secondary school-age girls. So, the project was not focused on re-enrolment as a key education outcome. Instead, the project focused on delivering training on life and business skills. In order to make sure that these interventions are effective and that the girls are learning, one condition to fulfil was to ensure that these girls attend until the end of the nine-month training cycles and graduate from the programme. In Cycle 1, 5,974 out of 8,450 (70%) of enrolled out-of-school girls attended the Safe Spaces until the end of the programme, and in Cycle 2, 8,060 out of 8,425 (95%) graduated. This achievement is connected to life skills and economic empowerment, which were the core objectives of the project.

Lessons learned

- **The implementation of an electronic monitoring system for attendance was not accompanied by an increase in attendance rates**. The effectiveness of an electronic attendance monitoring system was difficult to measure because the control group used very different measures of attendance. A comparison could only have been ensured if both treatment and control groups collected information in the same way and with the same regularity, which is very difficult in control schools that are located in educationally marginalised areas.

- **A two-hour tutoring session per week enabled out-of-school girls to successfully attend the whole nine-months educational cycle** delivered by Coca-Cola and to graduate from the programme. A more time-intensive approach may have conflicted with girls' other household duties and led to increased drop-out from the course.

3.4 What has worked, why and with what effects?

SPW projects aimed to address a range of barriers to girls being in school and learning effectively.

Our analysis covered barriers relating to: school; poverty; girls' aspirations and decision making; the community's attitudes towards education; and violence. To assess the effectiveness of projects' interventions, we used the mapping of interventions to present projects' findings about the effectiveness of their interventions[56] on the barriers identified in target and thematic areas.

---

[56] We indicate the origin of the findings by referring to individual Project Midline Evaluation Reports and/or Outcome Spreadsheets (for cases in which outcome data did not validate projects' claims). Where possible, we also triangulate projects' findings using the existing literature relating to intervention effectiveness for girls' education.

For a more detailed analysis of intervention effectiveness including summary tables for each intervention, refer to Annex E.

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

In this section, we report on the following barriers and thematic areas, in order of importance with regards to their impact on girls' education across the SPW:

- School-related factors;
- Poverty;
- Girls' aspirations and decision-making (early marriage);
- Negative attitudes towards girls' education; and
- Violence and safety.

3.4.1 School-related barriers

Key Findings

School-related factors are still prevalent barriers to education at endline. All three projects have designed and implemented interventions addressing school-related barriers. This mainly consisted of the provision of ICT equipment to schools (Discovery and Avanti) and teacher training by all projects.

Projects found evidence that due to their interventions, perceptions around teaching quality for Discovery (by caregivers), and the perception of having adequate teaching materials and learning aids have improved (Discovery and Avanti).

Projects also found that ICT equipment created a more interesting environment for teaching and learning, and improved teachers' confidence and the teaching and learning process. However, both Discovery and Avanti noted that teachers and schools faced serious challenges which limited the usefulness of the equipment and the learning content, and limited the anticipated effects on the education process and outcomes.

Our reanalysis of the project data found evidence that teaching practices did not necessarily improve across countries and across most practice areas. Yet, where they did improve, the evidence suggests that some teaching approaches were more suitable and therefore better for certain contexts over others.

What did projects find at baseline?

At baseline (See Annex E Table E.1), **school-related factors were reported as the most important barrier to girls' education** across the SPW. While school-related barriers were assumed by the three SPW projects at pre-baseline, Discovery (in all three countries targeted by the project: Ghana, Kenya and Nigeria) and Avanti provided evidence confirming the existence of these barriers, while Coca-Cola did not provide clear evidence to support this assumption.

Pathways through which girls' education is affected primarily relate to **poor quality of teaching; inadequate provision of teaching materials; inadequate provision of qualified teachers,** including teachers' absenteeism; unsafe **conditions for commuting** to school; and the **inadequate provision of schools and facilities,** which includes a lack of gender appropriate latrines.

Have interventions been designed and/or implemented to address the identified barriers?

All three projects designed and implemented interventions addressing school-related barriers to education (see Annex E- Table E.2). Discovery and Avanti focused on the provision of **ICT equipment as tools** to enhance teaching and learning in schools, accompanied by **skills training for teachers** (see Box 7 below). Discovery focused on all-classroom activity in which the students watch an educational video followed by the teacher leading an exercise, while Avanti focused on providing individualised support to students through the learning platform (accessed through the computers).

Discovery and Coca-Cola carried out interventions with community representatives to engage them in education. Discovery mainly worked with school-based management committees and PTAs. Coca-Cola focused on working

with **stakeholders at state, regional and community levels** to mobilise and encourage them to engage in girls' education and empowerment.

While **infrastructure and facilities**, including gender appropriate latrines, were found to be major barriers to education at baseline in at least two of the three projects (with Coca-Cola anticipating the barrier, but not reporting on it), none of the projects have designed or implemented interventions to directly address this. Additionally, while projects designed and implemented training for teachers, the issue of **teacher absenteeism** was also not directly addressed.

Box 7: School-related interventions by project

Discovery

Discovery's theory of change focused on core interventions relating to school factors – these are:

- **Provision of technology in classrooms:** 1470 schools established "Learning Centres" (LCs), which include: a large flat-screen television and DVD player; a second smaller TV and DVD player for teachers; and a library of educational DVDs.

- **Teacher training** provided to 15,383 teachers. The training focused on student-centred learning, activity-based learning, gender sensitive /responsive teaching practices, and the integration of materials from the LCs into teachers' daily curricula. Discovery's trainers did not provide all of the training to teachers (especially in Nigeria) due to the limited time available, the large number of schools involved and high level of teacher turnover. Instead, Discovery used resource teachers who are trained teachers to provide a shorter version of the training to their peers. According to project staff interviewed, teachers did not receive any compensation, but were rewarded with a certificate from Discovery and the Ministry of Education.

- **Mobilisation of parents, teachers and community members in education:** these activities were mainly aimed at producing Community Action Plans (CAPs), which include mapping of the assets available to schools, as well as identifying the main challenges to education, including girls' education. Templates and tools provided by Discovery were used for these workshops to produce these plans. Based on the CAPs, the community (through its representatives) engaged in advocacy activities to fulfil the actions in the plans and secure funding for them. According to project staff and the project completion report, committees managed to achieve some of these plans. For example, in Kano, the committees managed to facilitate building or refurbishing classrooms in 36 new girls' junior secondary schools.

Coca-Cola

Coca-Cola did not particularly focus on intervening directly within the school system. However, through their activities supporting girls in school, it carried out two main interventions addressing school-related factors. These were:

- **Training School-Based Management Committee (SBMC) members on leadership skills and resource mobilisation:** Coca-Cola worked with SBMCs, which include parents, Head Teachers, teachers, community leaders and members of civil society to promote girls' education. The project endline report mentioned some of the achievements of the SBMCs, which included fixing and refurbishing school facilities and equipment, and facilitating the provision of sanitary pads to girls.

- **Training for the Learning Space Coordinators (LSCs):** According to project staff, the LSCs recruited to support in-school girls were teachers, while the LSCs supporting out-of-school girls were volunteers from the community. The training for the LSC was between 3 and 5 days, but it did not seem to focus on teaching skills or pedagogy. For the LSCs supporting out-of-school girls, the training did not include specific academic literacy teaching skills.

Avanti

Avanti designed and implemented a number of interventions to address school-related barriers. These were:

- **Installing and maintaining ICT tools,** including connection to the internet and providing access to learning content. The project provided two types of ICT equipment: (1) computer labs (25 computers for an average of 1,000 students); and (2) projectors and laptops for teachers to use in class.

- **Providing learning content through iMlango platform** that included literacy training exercises for children, an African Storybook and a children's encyclopaedia. This content was approved by the MoE. It also provided the Maths-Whizz content focused on numeracy and maths. It is worth noting that these different content packages and activities were provided to different ages /grades, i.e. not all grades received all of the activities and packages. The project also provided in-field maintenance of equipment and support to teachers.

- **Provision of electronic attendance monitoring tool (sQuid cards)** to track and analyse students' attendance, as well as to help identify and address causes of student absence.

- **Training for the teachers** with support from the MoE and the Teachers' Service Commission (TSC). The training was for 3 to 5 days. Teachers were nominated by the schools and Head Teachers of the targeted schools. The project rolled out the training to other teachers, mainly English and maths teachers, but also other teachers.

What has changed since baseline?

At endline (See Annex E Table E.3), it seems that only Discovery was able to achieve some improvement in the barriers identified at baseline – most notably in the **quality of teaching**. There is also evidence that access to **teaching and learning aids** (material and resources) improved in schools supported by both Discovery and Avanti. This is not surprising as both projects directly provided teaching and learning aids to schools that teachers and students could use.

Despite Coca-Cola's engagement with **community and school management committees**, as well as **training Learning Space Coordinators**, there is no clear evidence in their evaluation report that school-related barriers to education improved.

Teaching and learning materials, resources and aids

Avanti and Discovery provided evidence that schools' access to **teaching and learning resources, materials and aids improved between baseline and endline**. This had a number of reported positive effects. Improved access to resources created a more interesting teaching and learning environment, encouraged attendance in schools, increased teachers' confidence and reduced their workload. However, a number of challenges were also reported as limiting the effectiveness of these resources and learning aids, which also limited their impact on the education outcomes that projects set out to deliver.

ICT equipment has created a more interesting environment for teaching and learning.

Both Avanti and Discovery reported evidence that the **provision of ICT equipment to facilitate and support teaching and learning created a more interesting environment at school** and generated interest around education more broadly – particularly for marginalised communities with limited access and exposure to technology and the 'outside world'. Students and teachers also had more resources to use in class.

According to Avanti's endline report, the majority of the targeted girls said that they found school more exciting because of the digital learning tools. Interestingly, the girls in the schools that only received projectors and laptops and not the computer labs showed the highest increase in improved perceptions about school. Our research, however, was unable to unpack and explain this finding.

There is some evidence that using ICT-based content improved teachers' confidence and the teaching and learning process.

Both Discovery and Avanti provided evidence that teaching improved due to the use of ICT equipment. The evidence was clear for Discovery and was described in a number of ways. For example, in Ghana, qualitative evidence described how the ICT tools and content reduced teachers' workloads to a more manageable level. This seemed to validate what the teacher was trying to convey, increasing teachers' confidence. In Nairobi and Nigeria, it was described in terms of helping teachers explain abstract concepts and therefore improving the students' understanding of the lessons.

"(Using LC) I have learned to lead the children to discover whatever they are supposed to know themselves without you necessarily have to do everything for the child to learn. The children are now prompted to the issues and it opens their minds to even think wider from what is just around them." (Teacher, Discovery, Ghana)

Similarly, Avanti provided some evidence that the online content made teaching easier – the endline report stated that 84.4% of teachers stated that digital learning tools made teaching easier – particularly when explaining abstract ideas; it also allowed teachers to get access to short videos and pictures, which they used for their own tailored teaching.

"*The teachers are browsing to get more materials from the internet. There is a day I went to* [a primary school] *and the teachers were busy browsing looking for more information to get more materials for teaching, so you find that it is very important."* (Community representative, Avanti, Kenya)

Students' use of ICT tools for learning relies on the teachers' engagement and confidence in using the ICT equipment.

Discovery and Avanti recognised that teachers needed to not only be able to operate the ICT equipment, but also to know how to integrate them into their teaching. Both projects provided training and support to teachers to enable them to effectively use the equipment (although according to interviews with staff, the extent of difficulty teachers faced in adopting the tools as teaching aids was not particularly anticipated by Avanti).

Specifically, Avanti found that the counties where students spent more time using the Maths-Whizz are where teachers were observed to be more engaged with their students. This suggests that where teachers are more engaged, students also are encouraged and guided to use ICT as learning aids.

While ICT equipment has created a more interesting teaching and learning environment and improved teachers' confidence, a number of issues limited its usefulness and effect on education outcomes.

There are a number of issues and challenges which project endline reports and project staff have reported to limit usefulness of the ICT equipment, and its anticipated effects on education outcomes.

Limited ability of teachers to operate and use the equipment. This was reported by Discovery and Avanti to sometimes be a challenge, particularly at the beginning of the engagement with the schools and later if trained teachers left the school. Discovery tried to overcome the issue by training resource teachers, who in turn trained and provided support to other teachers. Avanti on the other hand provided continuous field support and maintenance to schools. Yet, at endline, Avanti reported that more that 58% of the teachers did not feel that they were necessarily skilled in using computers and a quarter of the teachers in the treatment group did not use the digital tool that was provided. Additionally, the project noted that there were problems in applying the use of the Maths-Whizz, with the lessons starting from the beginning every time the student logged on instead of continuing from where the student was last time. The issue was later resolved by a system update. There is anecdotal evidence that Avanti (through its consortium member Camara) provided continuous in-field support to solve such issues. It is not clear to what extent this was effective or long lasting.

The equipment and rooms were sometimes crowded, inappropriate and not practical. Discovery's endline evaluation reported this issue primarily in terms of the number of students per Learning Centre (LC) classroom in all three countries. In Ghana, the LC class size ranged from 7 to 85 students. In Nigeria, the LC class size ranged from 8 to 120 students. In Nairobi, the LC class size ranged from 7 to 99, and most of these classes were taught by a single teacher. As this teacher explains:

"We have just a small room as a learning centre for a whole class so sometimes the teacher has to strategize to get the children in... It's consuming a lot of instructional time…." (Teacher, Discovery, Nigeria)

Additionally the physical state of the rooms was sometimes poor. In Nigeria, this was particularly described in terms of roof leakages. For Avanti this was also described in terms of the availability of computers for the students – on average there were 25 computers per school, often with 70 students in a class. Access to the computers was organised through timetables by class rather than for specific students. However, in their observations the evaluators found that students would rush to the lab, and access to the computers was, in effect, a first-come-first-served basis. The effectiveness of the equipment was also constrained by the availability of teachers to oversee students using the equipment, as well as their ability to use the projector and the computer labs simultaneously.

According to the endline evaluation report the projector would be kept in the computer lab for safety and transporting the projector to a room with electricity was a hassle, inefficient and sometimes impractical.

> *"We appreciate the idea, but we also want to request that if possible we need to have somebody on the ground, an ICT assistant for example. If one of your staff would be here permanently to keep the door open throughout the day it would expose these children more compared to a teacher who has something to do."* (Teacher, Avanti, Kenya)

System problems and a shortage of electricity/ fuel/ generators to operate the LCs (Discovery) and computer labs (Avanti) was a key challenge in all locations. In Avanti's case for example, the endline evaluation report states that computer-based learning was available for only 3.6 out of 6 hours during the school day, and in term three it dropped to 2.6 hours per day. This was particularly a challenge because the electricity bills for schools increased tremendously because of the electronics being used at school, which also exceeded the allowance allocated by the Ministry to pay electricity bills. In most cases, the school management committees were responsible for finding additional resources to pay the bills, but not all school committees managed to solve the issue. In general, these (sometimes lengthy) cuts in electricity resulted in disruptions in the use of the equipment.

There were also issues with the educational content and its alignment with the curriculum, context and lesson plans. While the educational content of Discovery and Avanti was approved by the relevant authorities, it was reported that teachers and Head Teachers sometimes felt that the content was not sufficiently aligned to the local context, school curriculum and lesson plans. In Ghana, some teachers in Discovery's schools expressed concerns about a lack of local focus, such as local history. Or as a Head Teacher in Wajir stated:

> *"Fear is that the content of the cassettes might be contrary to the cultural practices. It might be a taboo and it might cause shock. And you know when a teacher is somebody not from this community they might take it to be very positive for themselves only to see after they have started the lesson that it has caused a problem in the class."* (Head Teacher, Discovery, Wajir).

This was also found in Avanti's targeted areas in which some teachers and Head Teachers felt that the content was not relevant to the local context and culture, and did not see the added value of some of the content, which focused on other geographic areas (such as Asia and Latin America). It is not clear however whether the students felt the same about increased exposure to other parts of the world.

There is some evidence that the length of the videos provided by Discovery did not fit within the lesson plans and the accent of the narrator (described as a Western accent) in the videos was difficult for students and teachers to understand.

Additionally, Avanti found that some teachers were leading revisions for the exams using alternative resources (more traditional content) to the online content. It is not clear however, whether this is due to the inability of teachers to use the tools and content in an effective and useful way, or whether there are limitations to the content's relevance to the national curriculum.

Technology for school management

In addition to providing ICT equipment to enhance teaching and learning, Avanti provided digital-based systems for schools to enhance their attendance monitoring. The purpose of the system as described in their theory of change (ToC) is to provide real-time data for schools and authorities to better understand the attendance patterns of children and their marginalisation factors. This, as noted, would allow the school and stakeholders to better tackle marginalisation factors and barriers to education, and ultimately improve attendance.

ICT-based monitoring system encouraged students to attend school in a timely manner, but there is insufficient evidence to suggest that it has been effective in improving student absenteeism.

There is mixed evidence of the extent to which the real-time data was being used to tackle issues of marginalisation and barriers to education as suggested in Avanti's ToC. While the endline evaluation report did not provide strong evidence to support this, project staff provided anecdotal evidence of this happening – for example, where teachers have identified repeated absence of students, investigated the matter and reported it to the relevant school committee. It is not clear, however, what the scale and added value of the new digitalised tools are in comparison to the traditional way of keeping attendance records.

However, the report does provide evidence that the electronic monitoring system encouraged students to attend school. As illustrated in the report, this was demonstrated by students who feared being reported to the authorities as having missed school or being late for school.

"Children usually attend school every day and the child tells you, "Mum I can't miss going to school because if I get late, the computer will show that I was not in school. It is not like in the olden days when children used to hide in the bushes, you know that if its 8 a.m., you know you have to swipe in the computer, so it's a must by 8 a.m. to be in class so that is something that makes them not to be late for school as they do it daily." (Caregiver, Avanti, Kenya)

Teacher training and support

Discovery and Avanti provided evidence that teacher training and support interventions have positively affected teachers' skills and the classroom environment. This emerged from a number of interventions which provided skills training in learning and student centred approaches, gender sensitive practices, and in integrating LCs or digital-based learning into lesson planning.

As previously mentioned, Coca-Cola trained teachers who are the Learning Space Coordinators for in-school girls on numeracy, literacy and other academic subjects. There is no clear evidence that this training affected teachers' skills.

Teacher training had some positive effects on teaching quality, but with limited improvements in their confidence, particularly in the use of ICT equipment and their capacity to encourage students' creativity and critical thinking.

Discovery in all three countries reported that teaching quality had improved at endline compared to baseline. This was mainly measured through classroom observation, which specifically aimed at measuring "teaching best practices" (see Box 8), but also through interviews and focus groups discussions. As reported, the improvement was specifically evident in the use of a variety of activities, group work, gender equitable language, and actively involving non-participating students. Teachers themselves reported improved teaching methods but also a more gender-sensitive approach to teaching. As this teacher in Nigeria states:

"The change is, before Fitila[57], hardly will you see students doing group discussions. The group discussion helps them in understanding things because they discuss within themselves. There are introverts and extroverts amongst them. Some of the students don't understand from the teacher's way of teaching, but if you group them and they talk or argue within themselves, it helps them a lot. This has a long lasting effect on the students." (Teacher, Discovery, Nigeria)

The effects of improved teaching skills were reported to have a positive effect on students' motivation, classroom engagement and interest in learning in general. However, it is worth noting that households did not report changes in the quality of teaching (see below and our reanalysis from project data).

Box 8: Discovery's Teaching Best Practices

Discovery adopted the framework on essential teaching practices developed by the Centre for Educational Effectiveness. The framework includes six categories of practices, each has 3 to 7 specific practices which can be observed and noted in class. Discovery's external evaluator selected 13 specific best practices and developed a classroom observation tool to measure these practices. These are:

- Does the teacher have a lesson plan?
- To what extent did you think the teacher met the objectives of the class?
- Towards the beginning of the class, does the teacher allow students to share what they already know or think about the topic?
- Did the teacher use a variety of activities in the lesson?
- Did the teacher clearly explain the expectations for collaboration (working together) to students?
- Does the teacher encourage students to think creatively to solve real-world problems?
- Did the teacher ask higher-order questions?
- Did the teacher use gender-equitable language (in a positive way) or non-gender-equitable language (in a negative way) throughout the lesson?

---

[57] Fitila is the Hausa name used for Discovery Project

- Did the teacher call on or actively try to involve a student who was not participating?
- Did the teacher have students work together in groups or pairs?
- Does teacher provide positive, encouraging feedback to boys and girls?
- Did the teacher ask for the students' opinions?
- Did the student activities seem to contribute to reaching a learning objective in the class (clearly aligned with what the students were supposed to learn that day)?

While the reports generally note that there has been an improvement in the quality of teaching, the teachers' encouragement of more sophisticated and challenging ways of critical thinking was limited and did not appear to have improved. Also, in some of Discovery's intervention areas, teachers' motivation and confidence was reported to have increased more in the control group than in the treatment group. For Avanti, teachers' confidence in using ICT was reported to be mixed, limiting their use of the ICT equipment.

Avanti and Discovery provided possible explanations for the limitations in the effectiveness of their teacher training activities.

Teacher relocations, redeployment (particularly between treatment and control groups), and retirement meant that trained teachers moved away from intervention schools. Avanti tried to provide continuous support to fill in the capacity gaps of teachers in schools. Discovery adopted the model of resource teachers who provided training and support to other teachers. The effect of these efforts, however, were limited as evidenced by the limited quality of their teaching.

During interviews, the challenge in Wajir (Kenya) was particularly noted. There is not a sufficient number of teachers who are from the region, and due to violence in the area, many teachers relocated and not many trained teachers went back to teach in Wajir. Again, resource teachers were used, but this still limited the quality of the training and ultimately, the quality of teaching.

Insufficient teacher training: Avanti's endline evaluation report makes an attempt to link the limited use of ICT to the skills training provided to teachers, which, as suggested, might not have been sufficient. The report and staff also noted the difficulty in getting teachers to use the ICT-based tools for teaching, which needed more in-depth and prolonged engagement with the teachers.

Discovery, and particularly Discovery's project in Nigeria reported that not all teachers were trained by Discovery because of the large number of teachers in a school combined with the limited timeframe for the training. The report states that about 50% of the teachers in Nigeria were trained by the resource teachers and so did not receive the full set of training activities.

Evidence from the EM's reanalysis of project data

As discussed in this section, improving the quality of teaching was a central goal of projects in the SPW. In addition to the teacher training components of all three projects, the ICT components were also intended to improve teaching quality by providing effective teaching aids. While teaching quality is notoriously difficult to measure, SPW project data offers a few opportunities to analyse impact in this area.

One rough metric of teaching quality collected across all windows of the GEC is parental perceptions of teaching quality. For Discovery and Coca-Cola, this data is presented in Figure 30. Avanti did not have this question in their household survey. One complication of interpreting these responses is that the survey question asks if teaching has improved recently, not how good teaching is currently. This means that if we take these results at face value, a large proportion of families already believed that teaching was improving in all countries at baseline. And even among the control group the portion of families who believe teaching is improving increased from baseline to endline for Discovery. On the other hand, the results show little difference between the change in treatment and control groups. Discovery's programme in Ghana is a possible exception, where the portion of families who said education was improving increased by 11 percentage points more in the treatment group than in the control group.

Figure 30: Perceptions of changes in teaching quality (Discovery and Coca-Cola)
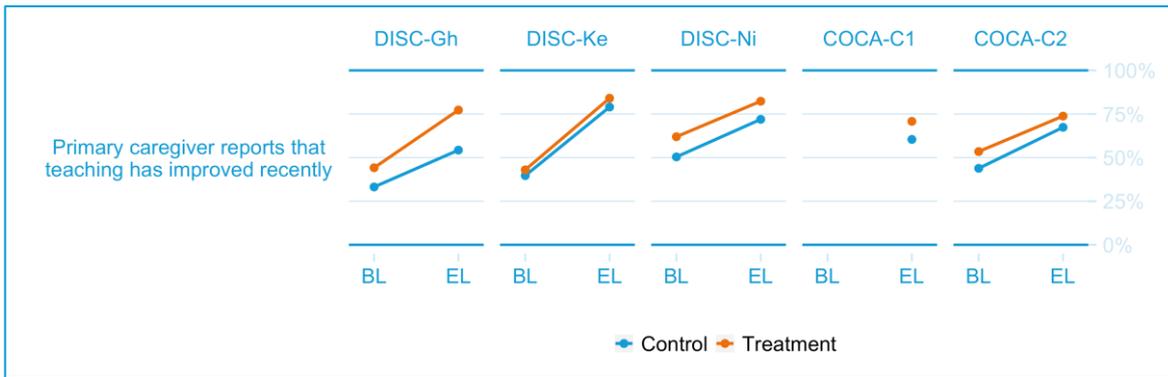
This finding in Ghana should be approached with some caution because there seems to have been a systematic pattern at baseline that caregivers were less optimistic about teaching in the control group than in the treatment group. We do not know what accounts for this difference, so it is possible that whatever caused this difference is also responsible for the larger increase in optimism in the treatment group over the course of the project. Still, if we take this together with other sources of evidence, there is a plausible case that Discovery's teacher training and ICT resources worked better in Ghana than elsewhere. In particular, at baseline, Ghana was the only country where the "student-centred" teaching methods seemed to have a positive association with student performance at the school level, and as we describe later in this report, caregivers in Ghana appear to be more engaged in their children's education than in the other countries. None of this is conclusive, but it does make a plausible case that in some of the communities where Discovery worked in Ghana, the teaching approaches they promoted were more suitable for the local context than elsewhere.

While parental perceptions of teaching quality are useful because we can compare this across all GEC windows, we have more detailed evidence of the changes in teaching quality in the classroom observations conducted by Discovery. While the sample size is relatively small (between 157-178 observations per country project) and subject to the consistency problems across observers, it is one of the larger and most informative datasets on teaching quality collected across GEC windows. In the baseline report for the SPW, we used the first round of classroom observations to explore the relationship between learning outcomes and teaching practices promoted by the project. In this analysis, we found a generally weak association between the use of best teaching practices identified by the project and learning outcomes, but we noted that in a few of the highest performing districts, particularly in Ghana, these practices were more common.

With two more waves of data collection (midline and endline),[58] we can build on this, by assessing how effective teacher training was at changing teaching practices. It is important to note that comparability of data at baseline with midline and endline is limited as it has been found that some data collectors were not sufficiently trained to identify Discovery best teacher practices. This issue has been corrected for the next data collection rounds. Discovery's endline evaluation report presents the difference in the average number of good practices used by teachers between treatment and control groups. It reports that the difference was 1.6 in Ghana, 3.3 in Kenya, and 1.3 in Nigeria. In

---

[58] SPW projects did not conduct a full wave of midline collection, but Discovery conducted a midline wave of classroom observations to inform implementation of their project.

Figure 31, we present 12 indicators of best teaching practices tracked from baseline to endline.[59] Inspection of these figures raises some questions about how training affected teacher behaviour. It would be expected that if training led to improvements in teacher behaviour, practices should improve in the treatment group and they should improve relative to the control group. But this isn't often the case. For example, for practices 3 and 4 in Kenya, asking students to share what they know about a topic and using a variety of activities, the main reason that treatment teachers did better was because this practice declined significantly in the control group. The actual increase in the use of these practices in the treatment group was small by comparison. In fact, Kenya is the only country where there is an overall improvement from baseline to endline in the number of practices used in the treatment group. In Ghana and Nigeria, most practices were used less at endline than at baseline in the treatment group.

The classroom observation tool can be somewhat subjective, and due to a lack of training of data collectors to identify these best practices at baseline, changes in the number of good practices observed over time are less reliable than changes in the difference between the treatment and control groups. For each wave, the same enumerators go to both treatment and control schools around the same time, so the comparison should be more direct than the comparison across waves. If this is the case, then there is some evidence that the teacher training led to improvements. In all countries, there were more practices where the treatment group improved relative to the control than practices where the treatment group declined relative to control. On the other hand, it appears that in all countries the treatment declined relative to the control group from baseline to midline, and most of the improvements were between midline and endline. It isn't clear why there would have been a decline from baseline to midline, which also suggests these measures may not be reliable.

Taking these results as a whole, it is reasonable to conclude that there is not a consistent trend that these practices are incrementally adopted more among teachers in the treatment group than in the control group in Ghana and Nigeria. In Kenya, there appears to a somewhat more consistent trend where several practices improved from baseline to endline and improved relative to the control group. However, some of the largest improvements relative to control still come from declines in the control group, so this finding should be taken with some caution.

---

[59] These are the teaching practices referenced in Discovery's endline evaluation report and discussed in Box 8. Within the classroom observation datasets, there are sometimes different variables measuring the same practice and different thresholds that can be set to determine if a teacher did exhibit that practice. In presenting this data, we are not trying to directly replicate the analysis conducted by Discovery, and our results may differ in some respects from the original analysis because of the choices made. Instead, the contribution of this analysis to look at how practices changed over time, and by using the same dimensions of teaching quality, we ensure that our analysis is compatible with the analysis conducted by Social Impact, Discovery's external evaluator.

Figure 31: Change in best teaching practices as defined by Discovery

Have changes in barriers had an effect on education outcomes (attendance, learning)?

As shown in Annex E table E.4, there is very limited evidence to show that school-related interventions have improved the education outcomes of girls. There is some qualitative evidence which suggests that digital tools and equipment, as well as improved teaching, encouraged students to attend school, particularly for Discovery in Kenya and Nigeria. Additionally, there is some weak quantitative evidence particularly in Ghana and Nairobi suggesting a link between improved teachers' practices and learning outcomes.

For Coca-Cola, there is anecdotal evidence from project staff noting that due to the training provided to teachers, scores of in-school girls in the targeted regions have improved in the national final secondary school examination. This, however, is not evidenced in the report.

Avanti used different intervention groups to identify whether each type of intervention had an effect on education outcomes. It found no evidence that its school-related interventions had a direct effect on education outcomes (assessed through statistical tests).

Additionally, the projects noted major limitations and challenges hindering the achievement of outcome results:

- Discovery's project staff noted that the focus of the design at the outset of the programme was not on numeracy and literacy, but on improving the life chances of girls. They also noted that in order to achieve significant rapid gains in literacy and numeracy, focused interventions are required that equip teachers to more effectively teach basic maths and reading skills.

- In the case of Coca-Cola and Avanti, the time allocated for teaching and learning numeracy and literacy skills is very limited. For example, Coca-Cola's in-school girls have about 2 hours per week in the Safe Spaces. The time is divided in the following: academic tutoring 50%, financial education 25%, and leadership and life skills (including vocational skills) 25%. This means that girls have less than an hour per week to train on numeracy and literacy skills, but also cover other academic subjects. Avanti has 25 computers per school with an average functioning time of 2.6 hours per day for the labs (term 3). It also provided evidence that the time on task per girl per week is not sufficient and doesn't meet the recommended time to achieve learning results.

3.4.2 Poverty

Key Findings

Poverty-related factors are still prevalent barriers to education at endline. Only Avanti designed and implemented a core activity (i.e. the provision of stipends) to address this barrier. Coca-Cola focused on income-generating activities and facilitated saving groups and other economic and financial-related interventions that were not intended to directly affect the barriers to an academic education that their target girls faced.

Projects reported that direct interventions to offset the cost of schooling increased attendance, while their indirect interventions such as Coca-Cola's income generating activities did not affect barriers to education or girls' education outcomes. Projects also reported that while there was some evidence that caregivers might have shifted their attitudes to understanding the benefits of girls' education, this did not translate into an actual decrease in girls' housework commitments that could have improved attendance and enrolment in school.

Our reanalysis of project data suggests that poverty remains one of the most important barriers to girls' educational achievement and that the projects had little impact in this area. Our reanalysis of Avanti's and Coca-Cola's data did not find any significant difference between the treatment and control groups with regards to the time girls spent on household duties, which remains a key barrier to girls accessing a quality education.

What was the situation at baseline?

At baseline (See Annex E table E.5) we found that poverty-related factors were ranked second among the main barriers to education. Poverty-related factors were anticipated by all projects, but Coca-Cola did not present strong evidence in its baseline report to support its assumptions about their effects on the success of the project's theory of change. Pathways through which girls' education were affected primarily related to extreme poverty and a lack of resources, which included material deprivation, limited resources and financial issues faced by families, forcing households to find and apply coping strategies in order to be able to send their girls to school.

Other poverty-related barriers described by Discovery relate to household commitments and income-generating activities and the cost of schooling that were reported by Avanti and Discovery. Avanti produced qualitative evidence that food and water shortages affected girls' education at baseline.

Have interventions been designed and/or implemented to address the identified barriers?

Out of the three SPW projects that identified poverty-related barriers to girls' education at baseline, only Avanti designed core activities related to economic interventions aiming at directly offsetting the cost of education and enabling girls to attend schools and learn from the outset of the project (see Annex E - Table E.6).

Coca-Cola trained out-of-school girls in income-generation activities and entrepreneurial, financial and business skills. The project formed savings groups and provided the target girls with assets in order to start their own income-generating projects. The design of these interventions, however, did not aim to offset the cost of schooling for girls, nor were they intended to enable girls to enrol in school.

Box 9: Projects' specific interventions

- **Discovery:** Discovery did not design any interventions targeting poverty although poverty-related barriers were found at baseline to be important. However, there is some evidence in their evaluation reports and from discussions with project staff that the girls' clubs, particularly in Nigeria, were engaged in income-generating activities, such as soap making, which supported girls' livelihoods.

- **Coca-Cola:** Coca-Cola designed income-generating activities for out-of-school girls (62% of the total number of girls reached by the project) through asset provision and creating saving groups for out-of-school girls (who are 18 years old and above), which aimed at improving the girls' economic situation. However, as the ToC set out and according to project staff, these interventions were not directly designed to offset the cost of education for girls, or to improve education outcomes for these girls and their families within the life of the project. Rather, Coca-Cola's intervention with out-of-school girls' was focused on improving their life chances by enabling them to increase their income and get more control over their lives through new life skills they acquired. However, from the interviews with the project's staff, it seems that the project was also trying to enrol out-of-school girls who were engaged in those activities in school and on other literacy courses.

- **Avanti:** Avanti provided stipends to increase attendance at school. At baseline, the project found that the most frequent reason given for girls dropping out of school was the inability to pay school fees. The project assumed that girls were not attending school because their families could not afford schooling. According to its endline report, Avanti identified students based on attendance criteria, and validated the final selection through a consultation process with the school and community committee to ensure that the girls were the most in need (see Section 3.1 on the project's reach for further information).

What has changed since baseline?

We acknowledge that it is beyond the capacity and mandate of the GEC to address structural barriers to education by alleviating poverty. However, we can investigate whether projects have been able to help girls and their families facilitate better access to education and quality learning by reallocating resources or bridging gaps in household finances. We can also measure whether families and caregivers decreased girls' household duties and their involvement in income-generating activities at home to allow them more time to attend school and do homework.

At endline, the available evidence from projects' reports suggest that only Avanti managed to have a positive effect on poverty-related barriers. Avanti reduced the cost of schooling as a barrier to education by providing stipends to some of its students. Coca-Cola reported that 'gatekeepers' (understood as caregivers and community leaders) reported a more equitable division of labour. Although poverty-related barriers have been observed at baseline, Discovery's programme does not address them, and the reports do not mention any change in these barriers at endline.

Cost of schooling

The available evidence shows that only Avanti was able to decrease the cost of schooling through the provision of stipends. While Coca-Cola's ToC includes income-generating activities, these were focused on out-of-school girls with no direct aim to offset the cost of their schooling, enrol them in schools or improve their numeracy and literacy skills.

Direct interventions offset cost of schooling and increased attendance.

Avanti provided stipends that were 'semi-conditional' on children's attendance in school. The stipends were non-conditional in part because the project believed it would have been difficult to apply fully conditional payments in practice. The stipends were disbursed through certain merchants who sold necessity items such as food, but also school-related materials.

According to Avanti's endline evaluation report, over 10,000 beneficiaries received about 1,000 KES[60]. To contextualize this, the report mentioned that a majority of parents (60.2%) spend 500-1,999 KES per year on education, which is a considerable proportion of their annual income. Therefore, we can assume that the stipends covered a substantial amount of their annual spend on education. The report also provides evidence that the stipends fulfilled their intended purpose because they were primarily spent on food, household items such as soap, and some school-related items such as learning materials and stationery.

While attendance was not entirely conditional to receiving the stipends, it appears that they had a positive effect on girls' attendance. The report states that more than half of the stipend recipients increased their attendance by any measure, and about 68% of the recipients increased their attendance by at least 10%. However, the report recognises that the provision of stipends is unlikely to have a durable effect.

While the direct provision of stipends appears to have a positive effect, there is no clear effect that income-generation activities offset the cost of schooling. According to Coca-Cola project's interviewees, the income-generating activities targeted out-of-school girls who were 18 years old and above (in accordance with local laws and legislation). The endline report also states that: 1,747 in-school girls (ISGs) were involved in saving groups activities; 7,473 (out-of-school girls (OSGs) were involved in saving groups activities; and 6276 OSGs who were 18 years old and above entered value chains (related to products associated with Coca-Cola and Light).

There have been some recorded successes, such as increasing the overall savings of girls in the savings groups; a reported 90% of Cycle 2 girls having increased their savings, and some girls have accessed loans to start their business:

> *"I have started a business from my savings selling pampers"*

> (Beneficiary girl, Coca-Cola Cycle 1, Lagos)

Despite the large number of girls supported, and despite evidence that girls increased their savings, **it is not clear to what extent the project has had a positive effect on girls' lives.** For example, the 2015 Annual Report for Coca-Cola reports that while girls appear to have increased their savings (some reported a maximum of N1300 in savings over nine months), the cost of the first crate of Coca-Cola (which was not provided by Coca-Cola) was reported to cost significantly more than these savings, at N1800. During interviews, it was mentioned that following advocacy on behalf of the girls, the crates were eventually provided to girls.

Another challenge which the project faced was the issue of training out-of-school girls below the age of 18 in income-generation activities. At the design stage, the project assumed that girls below the age of 18 could be involved in these types of activities, but this was prohibited by local laws and legislation. This left some girls frustrated at not being able to participate in these types of activities.

Additionally, there is no strong evidence that the interventions affected the girls' barriers to education, or affected their education outcomes (attendance and learning). There is however, anecdotal evidence, mainly from project staff, that out-of-school girls have become more interested in joining literacy classes and going back to school, although the latest project figures (in the project completion report) show that only 12 girls managed to re-enrol in school.

---

[60] On 1 July 2017, US$1 = 104 KES.

Access to finance for adolescent girls might not translate into increased spending on girls' education.

As discussed above, Avanti provided clear evidence that the direct provision of stipends allowed caregivers to address the cost of schooling as a key barrier to education. However, Avanti also stated as part of its qualitative data analysis of barriers, that once children grow older (i.e. reach adolescence), they are more likely to drop out of school to provide financial support for their families and because of the opportunity cost of going to school. If this is the case, then the size of a stipend would have to compensate for this lost income and not just the costs of schooling in order to prevent girls from dropping out.

Coca-Cola, which is also working with out-of-school adolescent girls, did not provide evidence to suggest that increasing the income of girls enables them to access education through increased spending on their education. Instead, the endline evaluation report states that "the importance of the financial stability of the gatekeepers cannot be overemphasized". It is worth noting that Coca-Cola could only provide income-generating activities for girls over 18 years old. This limits the effect of the project on younger girls.

Although a conclusion cannot be strictly drawn from these examples, the evidence might suggest that providing the opportunity for adolescent children to generate income is unlikely to lead to increased spending on their education.

Housework commitments

By the end of the GEC, we still find that housework commitments continue to be an important barrier to girls' education for all projects.

While attitudes of parents and caregivers might have improved towards girls' education, household chores for girls continue to be an important barrier to girls' education resulting in them not being in school.

Generally, projects tried to address this barrier through general awareness raising and sensitisation activities with parents and community members, stressing the importance of girls' education and trying to facilitate better access for girls to education. Households were encouraged to reallocate scarce resources and bridge gaps, suggesting a clear link between projects' interventions targeting this barrier and interventions targeting parental attitudes towards education.

While community mobilisation and sensitisation appears to have had a positive effect on parental or caregiver attitudes towards girls' household commitments for Coca-Cola Cycle 2, it does not seem to have worked for Discovery or for those girls participating in Coca-Cola's Cycle 1 course (see Annex E Table E.7). Discovery's endline qualitative data analysis flagged a relationship between household chores and girls' commitment to their education in Wajir, Nairobi and Nigeria. The project's reports found that while awareness raising has been carried out, and attitudes of parents and caregivers appears to have improved towards girls' education, this did not translate into a decrease in the amount of household chores girls had to do to allow more time to study at home. Evidence provided shows that compared to boys, girls are still primarily responsible for doing household chores.

Avanti reported no clear evidence of the extent to which the barrier exists, but found that the barrier nevertheless continues to exist. For example, the report discusses how girls in Kajado often have to assist their mothers in household duties and miss school. As described at baseline, there seems to be two contributing factors to this: poverty, and social and cultural norms. As illustrated by a caregiver in Kaduna:

> *"…her task is already defined by God"*

Evidence from the EM's reanalysis of project data

Reanalysis of project data provides some evidence to support the finding that poverty remains one of the most important barriers to girls' educational achievement and that the projects had little impact in this area. In our baseline report for the SPW, we looked for systematic differences in indicators of poverty between samples of girls who are in school and those that are out of school in the Coca-Cola and Discovery projects. Reviewing this evidence from our reanalysis of project baseline data, we concluded that no individual indicators of material deprivations were consistently associated with lower enrolment. However, in different contexts, particular material deprivations were more common among households with out-of-school girls. We also found that household perceptions of their poverty level may have a more important impact on their decisions to enrol their children than their actual living standards. Finally, we found in data from Avanti, that there was a relatively strong association between the amount of work that girls have to do outside of school and their reading test scores. Due to these results coming from cross-sectional data, this was only associative evidence.

Figure 32: Reasons given by primary caregiver for girl leaving school (Discovery, Ghana)
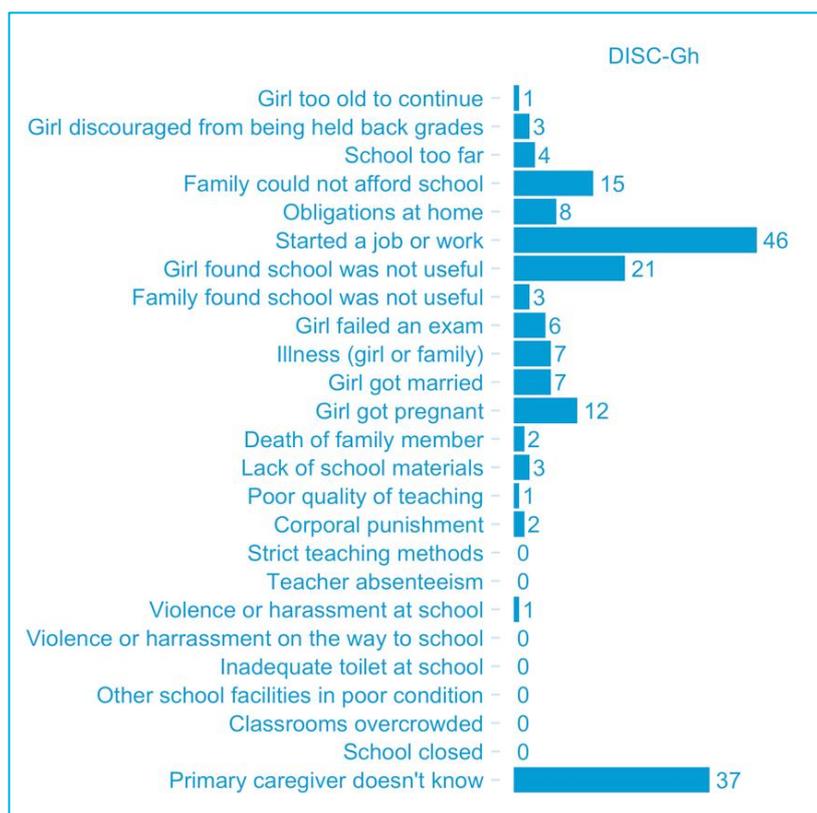
One source of evidence about the impact of poverty on enrolment at endline comes from Discovery's household survey, where the evaluator followed up with the households of girls who dropped out of school between baseline and endline. There are many advantages to this data source compared to the cross-sectional data at baseline. Due to the way the sampling for this project was done, it is reasonable to consider this to be a small but representative sample of girls who dropped out of school during this time period. The sampling approach was to select girls randomly in schools and then follow up with their family at their home. This meant that when there was attrition from the sample of girls selected in the school, the project could still follow up with the families of those girls to find out why they were not at school. For Discovery's project in Ghana, there was attrition of 421 girls out of a sample of 1978 and, of those girls, the project was able to follow up with 366 families. Of these, the primary caregiver of 186 girls reported they are no longer enrolled in school. With an 87% success rate following up with these families, the potential non-response biases among this group are relatively small. Figure 32 shows the frequency of different reasons given by the primary caregiver for why these girls dropped out.

These responses suggest that systemic poverty is the largest cause of girls dropping out of school. The most common reason given was that the girl had left school to work, with a significant number also reporting that they could not afford school or that the girl had too many obligations at home. If we take a broader view of how systematic poverty affects the decisions of households, many of the other reasons given for dropping out also fit with this hypothesis. A significant portion of families said they or the girl did not find school useful, which may be a reflection of the lack of opportunities they believe are there.

At first glance, this also appears to support a hypothesis put forward by many GEC projects that one of the major ways poverty affects educational outcomes is through the time girls have to commit to work and other tasks outside of school. Since taking work was the largest reason girls dropped out of school, it would seem likely that these kinds of commitments affect girls' attendance and diligence while they are in school. However, the data from Discovery, Ghana does not support this.

Figure 33 shows the distribution of the portion of time girls spent on duties outside of school comparing the girls who did drop out to those who did not. Surprisingly, there is no difference between the groups. On average, girls who dropped out were reported to spend 23% of their time on duties outside schools, compared to 26% among girls who did not. Even comparing the girls that dropped out to start a job or work, this group only reported spending 33% of their time on duties outside of school on average. As shown in Figure 33 there also was little

difference in the distribution of time girls spent between the two groups: there was not a larger cluster of girls who spent most of their time working or in any other way to distinguish the two groups.

Figure 33: Histograms of time girls spend on duties comparing girls who dropped out to girls who did not at baseline (Discovery, Ghana)[6162]
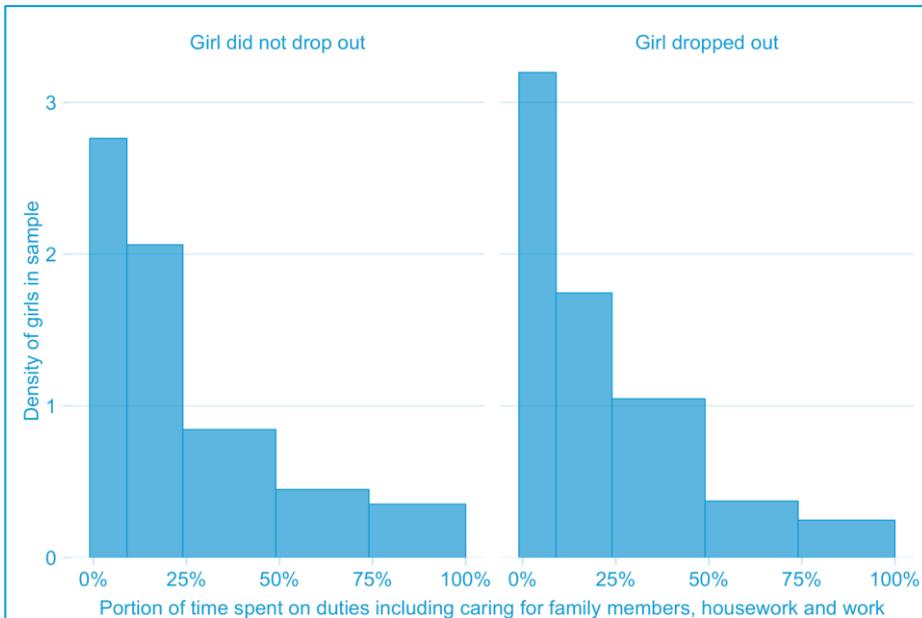
Another notable aspect of the responses given in Figure 33 is that very few responses related directly to the quality of schools or teaching. While a significant portion of respondents said that the girl found that school was not "useful," it is not clear whether this was because the school was poor quality or because the girl and/or her family felt that the skills learned at school would not be helpful. However, since almost no respondents gave reasons directly related to the quality of schooling, it appears that at some level poverty is a more fundamental determinant of a family's decision to keep their child in school or not.

Unfortunately, the follow-up interviews with the families of girls who were not present at school had a lower success rate in Kenya and Nigeria, so we cannot compare these results across project areas. For these projects sample sizes are too small and there is too much room for analysis.

Despite the importance of poverty, we would not expect to see a direct impact of SPW projects on the overall incidence or depth of poverty within the timeframe of these projects. Poverty is a systemic problem in the communities where SPW projects operated, and none of the projects attempted to directly reduce the level of poverty. However, following the projects' theories of change and the evidence in their endline reports, projects could have impacted on barriers related to poverty by reducing the time girls spent on work and duties, or helping families afford the cost of schooling. Looking for these effects, it is important to bear in mind that the interventions directly affected the financial situation of families – Avanti's stipends, Coca-Cola's support for savings groups – were minor components of the projects. Other programme components aimed at changing attitudes toward girls' education could have had an effect on girls' duties, but we would also expect these effects to be small. It is not surprising that we do not find that there was any significant difference between treatment and control groups in the time girls spent on household duties.

---

[61] The y-axis on this figure is the density of girls in the sample, which means that it is scaled so that the area of each bar is a number of between 0 and 1 and the sum of all these areas is 1. The area of each bar represents the percentage of girls who work that amount of time on the x-axis. By inspecting the shape of each histogram, this gives us a visual way of comparing the distribution of girls' duties across the population. This rescaling also makes it possible to compare the distribution of duties between girls who did and didn't drop out directly even though the sample size of girls who did drop out is much smaller.
[62] On both graphs, the first bar refers to 0% to 10% range of time spent on duties.

Figure 34: Change in indicators related to girls' household duties and time spent on work
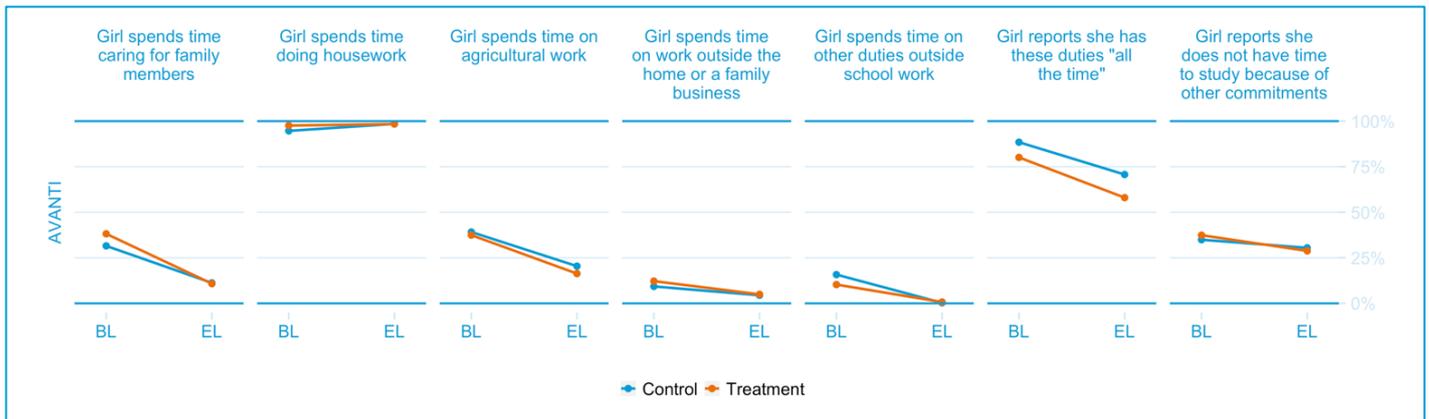

Figure 34 shows the change in several indicators related to girls' household duties and work for Avanti's project. It is clear from this evidence that there is no significant difference between the changes in the treatment and control groups. Given the approach Avanti took in attempting to target girls who are poor and do not have a good attendance record for stipends, this project would have had the best chance among SPW projects of affecting this indicator through financial support. One other notable feature of the evidence from Avanti is that there was a general trend across treatment and control for girls to report having fewer duties and spending less time on their duties, but there was no change in the last indicator in Figure 34: 'the girl feels her duties prevented her from studying'. It is possible that the decreases in duties are mostly an artefact of changes in the way the survey was administered between baseline and endline. However, if we accept this data, it suggests that marginal changes in duties do not necessarily have a direct effect on how much they can concentrate on school.

Figure 35: Family perceptions of hardship of school expenses
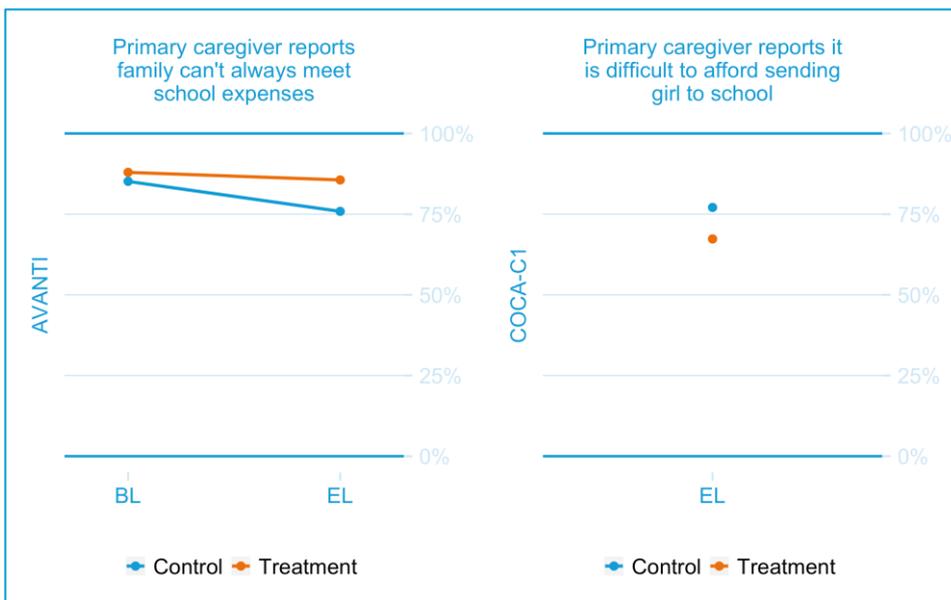
Limited evidence was available to assess whether Avanti's stipends and Coca-Cola's savings groups helped families to afford school. Figure 35 presents the responses to simple questions to the girls' primary caregivers asking whether they can afford to send the girl to school. Avanti's data shows that the situation changed little for either the treatment or control groups.

Have changes in barriers had an effect on education outcomes (attendance, learning)?

Only Avanti observed a positive effect of its intervention on education outcomes, and this was mainly on attendance. However, this was strictly observed on the students that received the stipends and did not reflect an overall increase in attendance at the project level. Please refer to Section 3.2 (Outcome section), for further discussion based on our analysis of education outcomes and attribution to projects' interventions.

3.4.3 Girls' aspirations, decision making and early marriage

Key Findings

There is some evidence that girls' aspirations and interest in schools has improved from baseline to endline. This is due to a number of factors including the effects of girls' clubs and safe spaces. Girls' clubs and safe spaces in particular appear to have improved girls' wellbeing and confidence, but girls still face challenges related to decision-making around schooling and marriage.

There is some evidence to suggest that girls' clubs with a focus on tutoring affect girls' learning. This was observed for both Avanti (through qualitative evidence) and Discovery (through qualitative and also quantitative evidence). However, both these findings need to be treated with caution because there was no evidence that these changes were attributable to the project and because these findings cannot be generalised – for example, there is evidence to suggest that Discovery focused on high achieving girls.

What was the situation at baseline?

At baseline, projects found several barriers related to girls' aspirations and decision-making around education, in addition to incidents of early marriage and pregnancies (see Annex E- Table E.9). Early marriage and pregnancy were described as an alternative to education in some contexts, and their prevalence was found to be strongly correlated with poverty levels. Most projects in their targeted countries (except for Discovery's project in Ghana) found evidence that early marriage is a barrier to education. A lack of female role models was assumed as a key barrier by Discovery, but was not directly assessed. Discovery's project in Kenya and Coca-Cola found other education barriers related to girls' aspirations and decision-making, such as older girls' lack of interest and motivation towards certain subjects such as maths and science and their increased interest in boys affecting their schooling.

Have interventions been designed and/or implemented to address the identified barriers?

Projects aimed at addressing the barriers to education related to girls' aspirations, decision-making and early marriage through two main categories of interventions: (1) extra-curricular activity and non-formal education; and (2) interventions to improve empowerment and self-esteem (see Annex E- Table E.10). Discovery and Coca-Cola designed and implemented core interventions to specifically and directly affect these barriers to education. For example, Discovery included girls' clubs with the explicit aim of improving girls' self-esteem and aspirations. Coca-Cola also included specific sessions in its education cycle aimed at developing girls' life skills, leadership skills and decision-making. Avanti's ToC however did not focus on this type of barrier to education (i.e. it was not a core intervention). Avanti instead intended to affect this type of barrier indirectly through: the girls' clubs and mixed clubs; some of its learning content; and the provision of ICT equipment to increase motivation and interest in education and increase exposure to the outside world.

Box 10: Projects' specific interventions

All three projects designed and delivered interventions related to extra-curricular and non-formal education activities, and girls' empowerment and self-esteem:

- **Discovery:** The project formed 927 girls' clubs. These were spaces where girls can choose to discuss or carry out activities as they desire. Discovery only provided toolkits for facilitating discussions in those clubs but adopted a flexible approach with regards to the actual content discussed in these clubs. Girls engaged in activities including cooking, soap-making, sewing, and gardening. These were also considered as safe spaces where girls could discuss topics such as education, hygiene, life at home, challenges faced as a girl, HIV/AIDS, peer pressure, careers, menstruation, discrimination, family issues, female circumcision, sex, early marriage, and pregnancy. These clubs and their activities intended to enhance girls' aspirations, motivation and confidence.

In addition to the girls' clubs, Discovery also produced and broadcasted National Talk Shows in each of the countries. These talk shows, as described in the project proposal, aimed at presenting role models for parents and girls to inspire them and encourage them to seek education.

- **Coca-Cola:** Coca-Cola introduced the safe spaces in schools (for in-school girls which comprise 38% of the total girls reached) and outside of schools (62% of the total girls reached). Coca-Cola adopted a structured approach where specific topics were discussed. Girls from the two groups were also trained in life skills and leadership skills to increase their self-esteem, confidence and negotiation skills through the Life Skills curriculum. According to project staff, Coca-Cola also created opportunities for girls to practice leadership skills, such as through the 'girls' ambassadors' model. Through this model, girl mentors are identified within the communities, and provide support to other girls and follow up on their savings.

  Coca-Cola intended to pair girls with other micro-enterprise owners to overcome obstacles to free movement to markets as well as to troubleshoot their problems and increase their confidence. This however did not materialise during implementation. During interviews, project staff noted that this was due to the limited timeframe and resources available. Instead, emphasis was placed on the girls' ambassadors model.

- **Avanti:** Avanti's ToC did not include interventions designed entirely to address girls' aspirations and motivations. However, through its ICT content, Avanti did aim to raise self-esteem, career aspirations and knowledge of the outside world. The project design also suggested that the provision of ICT equipment and content would increase interest in education and motivation to attend school.

  Avanti included child clubs as part of its ToC to enhance girls' self-esteem, and intended to deliver content such as the Good School Toolkit and the Q-files online encyclopaedia. The project completion report states that 387 child clubs (209 of which are girl-only) were established.

What has changed since baseline?

There is some evidence that **motivation and interest in school or certain subjects has improved for girls**. This was found for both Avanti and Discovery (in three of its four locations). Discovery reported some evidence that girls are more interested in science (particularly in Nigeria and Nairobi) and maths (particularly in Ghana and Nairobi). At baseline, the subjects of science and maths in particular were described in Kenya as male subjects. However, the endline report in Nairobi shows that girls from both treatment and control groups found school interesting and that more girls in the control group were eager to go to school than in the treatment group (97% in the control group compared to 91% in the treatment group). Avanti reported some evidence that there was an increase in the interest in school in general.

Unpacking the pathways through which changes have occurred, both Avanti's and Discovery's ToCs make links between girls' motivation and interest in school and the introduction of ICT equipment, teacher training (Discussed in school-related factors), but also Girls' Clubs. The latter is discussed below.

The evidence is mixed across projects with regards to changes in **girls' self-esteem, aspirations and their decision making power**. In both of Coca-Cola's cycles there was evidence that girls' confidence at school and at home improved at endline compared to baseline in the treatment group compared to the control group. However, there is inconclusive evidence with regards to changes in girls' decision-making power. The evaluation report for Cycle 1 reported evidence that there is some autonomy in girls' decision-making over resources, but decision-making about marriage is still within gatekeepers' hands. Evidence from Discovery and Avanti about these types of barriers is either inconclusive or shows that these barriers still exist, suggesting that there was no major change in its prevalence. Avanti, for example, presented positive evidence with regards to improved perceptions among girls towards early marriage and decision-making. It appears that girls are still not able to discuss these issues with their parents.

With regards to **early marriage and pregnancy**, none of the projects managed to demonstrate robust evidence that these barriers to education have changed. Similarly, only Discovery's project in Ghana was able to provide some evidence that the there was an improvement in terms of **girls having role models** who inspire them and encourage them to pursue their education.

Girls' Clubs

Extra-curricular activities, life skills clubs – whether open to girls only, or also to boys – were commonly established as a way to help students develop non-schooling skills, develop greater aspirations, increase their interest in school and education and their commitment to study. They can also build self-confidence and communication skills, which enable girls to participate more fully in class.

Despite the numerous clubs established, there appear to be some challenges in creating buy-in from parents and community members, and in their operation.

As mentioned, all three projects managed to establish and operate girls' clubs or safe spaces. However, Avanti and Discovery reported challenges that they faced during implementation. They did not specify how these affected the anticipated results of the clubs. The challenges that Discovery faced were mainly reported in Ghana and described in terms of limited support from parents and teachers for these clubs as well as limited financial support to run them. It is not clear to what extent these challenges affected the operation of these activities.

Another concern with regards to Discovery's clubs was the recruitment of girls in the clubs in which girls are selected based on their motivation and commitment to education. One teacher in Nigeria described the process:

> *"It's not done randomly. While teaching in class, there are students that put in great effort while lessons are in progress. They take their work seriously and they are also supported at home. These are the types of students we recruit."*

It is not clear to what extent this recruitment process is followed but it might suggest that many girls who are already educationally marginalised (and possibly generally marginalised) are unintentionally excluded from these clubs.

There was also evidence that Avanti did not particularly focus on this activity during implementation. This was mentioned during interviews and in the endline report. Clubs were optional for teachers and they were not provided with incentives to compensate for the extra efforts they put into the clubs, which limited their commitment to them. Additionally, one interviewee noted that project staff did not particularly focus on the clubs' interventions. It was also mentioned that Discovery's model of girls' clubs would be replicated in Avanti's schools in the next phase of the project, and that these would be delivered with the support of Discovery.

Girls' clubs and girls' safe spaces appear to be effective in increasing girls' wellbeing and confidence, although there is limited evidence demonstrating their effects on education.

Despite the challenges and limitations that were experienced, projects presented evidence that girls' clubs and girls' safe spaces had a positive effect on girls' wellbeing and confidence. For example, Coca-Cola found that safe spaces helped girls 'reason better' and voice their opinion without fear of retribution or rejection. Avanti's endline report described the benefits in terms of increasing their skills, including learning new things and improving their grades. Discovery's project in Ghana, Nairobi and Wajir also found positive effects on girls' wellbeing and confidence as a result of these clubs.

However, there is insufficient evidence to suggest that increased confidence and skills enabled girls to overcome the barriers to education related to their interest in school or their power to make important decisions about going to school. In Nigeria for example, more girls in the control group stated that they had decision-making power regarding schooling than in the treatment group. However, in Ghana, while girls in the treatment group were more involved in girls' clubs, there does not seem to be a difference between the treatment and control groups with regards to girls' eagerness to go to school. This suggests that while the clubs might have had a positive effect on girls' wellbeing (ranging from personal hygiene to self-esteem), this is not sufficient to affect barriers to education.

While safe spaces and clubs aimed to indirectly address early marriage and pregnancy, these remain key issues and barriers to girls' education that did not change as a result of projects' interventions.

Despite projects attempting to indirectly address issues of early marriage and pregnancy in their clubs and safe spaces, and despite the fact that there are indications that more girls are in favour of delaying marriage, these remain important challenges to girls and their education. This was evidenced in a number of ways, particularly in terms of girls' inability to make a decision regarding marriage. For example, one girl in Kano, Nigeria stated:

> *"If my father tells my mother to stop me, I will have to stop going to the school."*

Additionally, the prevalence of early marriage and its effect on education was demonstrated when parents and teachers were asked about the main reason for girls dropping out of school at endline. In Ghana for example, early marriage was still the most important reason for girls dropping out of school.

Cultural practices and social norms with regards to early marriage in particular are recognised by projects to be a difficult and persistent barrier that requires more than just empowering girls themselves, but also ensuring an enabling environment for change in these practices. For example, while noticing more interest in school and improvements in self-esteem, Discovery also found that parents are still the dominant decision-makers when it came to their daughter's future.

Have changes in barriers had an effect on education outcomes (attendance, learning)?

The evidence around the effects of most extra-curricular activities, non-formal education and self-esteem / empowerment related interventions has been inconclusive or ineffective with regards to improving education outcomes (see Annex E- Table E.12). There is some indication, however, that **girls' clubs with a focus on tutoring affected girls' learning.** For example, although Avanti did not focus on girls' clubs, a majority of girls stated that the clubs helped them improve their grades. Many also indicated that their favourite activity during the girls' club was the Maths-Whizz content. This might suggest that the girls use their time in the clubs to do maths exercises. This was also supported by the project staff during interviews who mentioned that students use the weekends and time after school to access the online learning tools.

Across Discovery's three target countries, there is some evidence that participation in girls' clubs (although some the evidence relates to both treatment and control schools) was associated with higher maths and reading scores. This finding, however, needs to be treated with caution because of the way girls were recruited in girls' clubs, i.e. teachers identified girls who were engaging well in class to participate in the clubs. This might suggest that the more educationally marginalised girls or those who were less interested in school were not recruited.

3.4.4 Negative attitudes towards girls' education

Key Findings

There is some evidence that Discovery managed to increase parents' engagement in children's education. However, there is mixed evidence suggesting that community outreach and awareness-raising activities carried out by Discovery and Coca-Cola improved attitudes towards girls' education. The evidence also suggests that parents' engagement was limited to those parents already interested in education, 'better-off' households, and those qualified or educated to a certain level. The reported evidence suggests that the national TV talk shows reached a very small number of the targeted communities.

Our reanalysis of the projects' data found that parents typically report positive views about girls' education. Our analysis also suggests that approaches to changing attitudes that are not targeted at particular individuals (such as community awareness activities) or focused on specific issues will not be successful. Since the members of the communities believe at a minimum that it is important to say that girls' education is important, any meaningful effort to change attitudes would need to identify a more specific problem relating to attitudes.

What have projects found at baseline?

At baseline, (see Annex E - Table E.13) the SPW projects anticipated barriers related to communities' and caregivers' attitudes towards education. Discovery and Avanti found evidence in their targeted communities that parents and caregivers have negative attitudes towards girls' education and that they do not support girls' education. Discovery described the barrier especially in terms of parents under-valuing investment in girls' education, as girls would eventually get married and leave their families behind, or in terms of favouring boys' education over girls' education. Coca-Cola on the other hand, assumed that parents and caregivers had negative attitudes towards non-religious education and post-primary education, but did not report on their findings at baseline. Avanti assumed that parents and caregivers perceived girls' education as irrelevant but did not find evidence of this.

Have interventions been designed and/ or implemented to address the identified barriers?

As Annex E- Table E.14 shows, Coca-Cola and Discovery have both designed and implemented interventions relating to parents' attitudes and/ or engagement in their children's education. Discovery stated in its proposal that it aimed to engage Community Action Mobilisers to work with each of the Learning Centres to set up or energise parent-teacher community associations and develop Community Action Plans to engage parents and the community in children's education. Discovery also produced and broadcast national TV shows in each of its targeted countries (26 episodes) aimed at discussing issues related to girls' education (including discussing the value of education). These were intended to shift the mind-set toward girls' education by improving parents' knowledge, attitudes and practices.

Coca-Cola carried out sensitisation and advocacy activities with the communities as well as national stakeholders in order to introduce the project and its value to girls and their communities. According to project staff, the project's local partner "Girl Effect" established and worked with Community Action Committees, which are comprised of community leaders and members such as traditional leaders, youth, women and the local government for the out-of-school girls' intervention, and School-Based Management Committees specifically for the in-school girls' intervention. It formed State Advisory Groups for the project to ensure local authority support, and implemented Champion Days to engage families and communities in girls education. At national level, the project also engaged with the media, other development organisations and the government, including by establishing partnerships with the National Youth Service Corps to provide graduate volunteers to support the programme[63].

Avanti did not design specific activities to address parents' negative attitudes towards girls' education or to increase their support for education.

What has changed since baseline?

At endline, and as Annex 5- Table E.15 shows**,** there is some evidence that Discovery managed to increase **parents' engagement** in children's education due to its mobilisation activities. However, there is mixed evidence about whether Discovery and Coca-Cola improved communities' **attitudes towards girls' education** or improved their perceptions of the value of their education.

There is mixed evidence with regards to the effectiveness of the community outreach and awareness-raising activities in improving attitudes and support for girls' education.

There is mixed evidence with regards to the effectiveness of activities aiming to improve communities' perceptions about girls' education. Coca-Cola, for example, carried out activities that target communities' perceptions about girls' education, including girls' knowledge and skills on vocational training and non-academic subjects. The Coca-Cola Cycle 2 report noted that more parents and community members perceived that girls' education and learning new skills is important at endline compared to baseline. However, both cycles provided evidence that this was not the case in Kano where attitudes towards girls' education do not seem to have positively changed. To illustrate this, it appears that in Kano parents still prefer to send boys to school and not girls. Also, while it seems that attitudes towards girls' education might have improved, the gendered roles and responsibilities have not, placing even more weight on the girls to support their families and community.

> *"…So it is very good for a girl to be educated because she will be a mother tomorrow, right?"*
> (Caregiver, Coca-Cola Cycle 2, Nigeria).

Similarly, Discovery also found mixed evidence in terms of the effectiveness of the community gender sensitisation (through the Community Action Plans) in changing attitudes towards girls' education. In Ghana for example, while caregivers themselves expressed support for girls' education, they noted that it was uncommon in their communities to send girls to school and that their community did not value education. In Nairobi, community awareness raising is perceived to have positively affected girls' attendance in school due to an increase in parents' interest in education and their encouragement of their daughters to go to school. However, some expressed their preference to send boys to school if they had to choose. In Nigeria, the endline report provides evidence that parents perceived that the value of education declines as the girl gets older.

Avanti did not particularly address the issue of communities' attitudes towards girls' education, but it provided mixed evidence of how parents and caregivers perceive girls' education.

---

[63] Project staff interviews and Project Management Response to the external evaluation.

There was little available evidence to explain why changing attitudes and practices towards girls' education was difficult to achieve despite the projects' interventions. During interviews with project staff, one interviewee from Coca-Cola mentioned that this might be due to the depth of the social norms and gendered roles within society, in particular girls' destined role in marriage and raising children, and little perceived value in girls' education in comparison to boys' education. As explained, changing these social norms and practices, particularly under difficult circumstances such as the financial constraints in which parents have to constantly make choices, is challenging. This is also mentioned in Discovery Wajir's report that explains how parents still prioritise boys' education when facing severe financial constraints.

The evidence suggests that the parents engaged by projects were already interested in education, from 'better-off' households, and were qualified or educated to a certain level.

There was also evidence that efforts to engage parents in children's education have had limited results. For example, in Ghana it is reported that more caregivers visit schools, but in Nigeria the report states that the parents' visits to school remained the same between baseline and endline. In Nairobi, while caregivers are reported to have visited the school, there was no change in the parents' involvement in school activities. It was difficult to explain why projects found it difficult to engage parents in children's education. Suggested explanations were illustrated in endline reports and during interviews with project staff, such as it is easier to engage parents of children who are performing well in school than it is to engage parents of children who are not performing well because they do not have high aspirations for their children's schooling. As this caregiver described:

> *"Most of the time you will find that the parents who attend the meetings are the ones their children perform very well. Some parents assume that there is no need to go for those meetings because their children do not do well in school. They do not see the need."* (Teacher, Discovery, Nairobi (Kenya))

One interviewee suggested that it is difficult to engage parents who are struggling to make ends meet. This is because these parents and caregivers do not have the time to participate in school activities. The interviewee also mentioned that in Kenya in particular, and since 2014, the government has changed the requirements of who can participate in school management committees and restricted them to those with secondary education.

There is no evidence that mass media reached the targeted communities, or that it had an effect on the targeted communities' attitudes and support for girls' education.

Discovery's second main intervention was aimed at changing communities' perceptions towards girls' education through national chat shows. As previously mentioned, these cost about 13% of DFID's matched funding to Discovery's project. They were presented in the English language and produced and broadcasted locally in each of the three countries. In Northern Nigeria, Discovery adapted the content and language of the broadcasts to reach the Hausa speaking population post midline, the programming was then broadcast widely in 2016.

Discovery's evaluation reports revealed that these shows did not have an effect on the targeted communities' perceptions towards girls' education[64]. This was mainly because most of the targeted communities did not watch the shows in their respective countries and on their local channels. 2.4% of households surveyed in Discovery's project areas reported having seen the programme (1%, 1.4%, and 4.5% for Ghana, Nigeria, and Nairobi respectively).

The low reach of the programme was due to limited access to televisions (especially in Ghana and Nigeria) and the language barrier (especially in Nigeria). In Nairobi where the viewership was the highest in project areas, focus groups revealed that the time of the programme conflicted with soap operas so viewers would switch between channels. It is also worth mentioning that the attitudes towards the importance of girls' education of respondents who have viewed the shows were positive both before and after the shows were rolled out. It is also worth noting that Discovery produced the National Talk Show in Hausa language and aired it on a Nigerian channel that broadcasts for Northern Nigeria - however, this was after the endline data collection.

---

[64] The external evaluator of Discovery carried out separate studies in urban areas using focus group discussions to assess the effect of the national talk shows on people's perceptions of girls' education. These studies reveal that the show was positively perceived and had some positive effect on encouraging and empowering girls, particularly through education. However, the purpose of these studies, the design of the tool and the recruitment of participants do not necessarily fit within the scope and purpose of the GEC evaluation framework. For example, the focus groups for Ghana were carried out in Accra, where the project did not intervene, which means that the possible effects would not impact the project's targeted girls.

Evidence from the EM's reanalysis of project data

In principle, it should be easier to measure negative attitudes and changes in attitudes over time using survey data. Surveys are best suited to gauge perceptions rather than constructs such as teaching quality. However, survey responses across all GEC windows have found less direct evidence of negative attitudes than we might have expected. Figure 36 presents the responses of the primary caregivers to a series of questions about their views of the value of education.[65] The most notable thing about the responses to these questions is just how unusual it is for parents to report a view that could mean they do not value education. The exception in the SPW is Discovery Nigeria, where between 30% and 50% of parents are willing to say they think it would be better for their daughter to work or get married at age 18 rather than continuing their education.

**Figure 36: Primary caregiver's perceptions of the value of education**
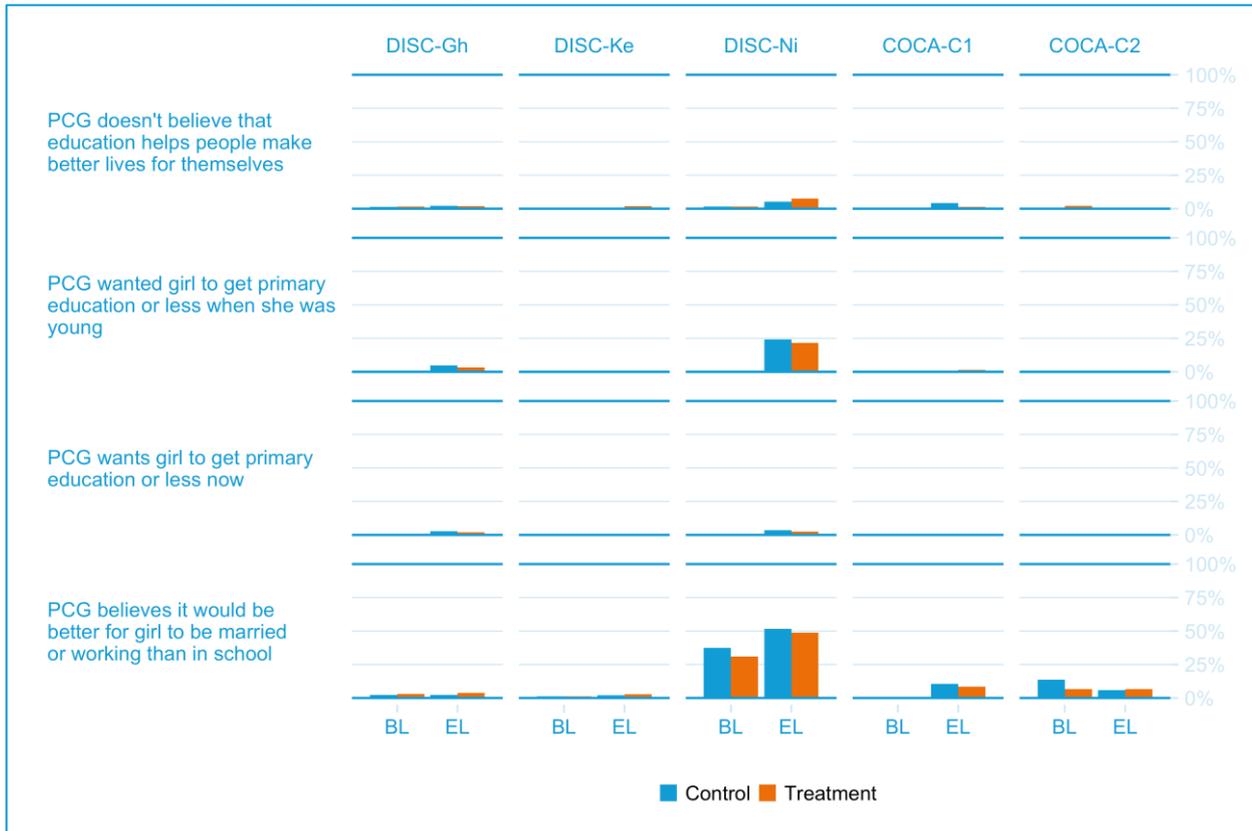
Taken at face value, these responses seem unrealistically enthusiastic about education. World Bank data on tertiary education does not have good coverage over all three countries, but suggests that female gross enrolment in tertiary education is between 5% and 15% in these countries.[66] Yet apart from Nigeria, nearly 100% of parents report they want their daughter to continue to tertiary education. In interpreting these responses then, we can either believe that parents are irrationally optimistic about their own children's educational prospects or that they are trying to project a positive attitude toward education even though they may be more realistic privately. But in either case, it is hard to reconcile these responses with the hypothesis that negative attitudes among girls' parents are a major barrier for girls getting an education.

The picture looks somewhat different when we look at how caregivers view the attitudes of the community more broadly. Figure 37 shows caregivers' responses to two questions related to support in the community. While caregivers themselves are unlikely to report views that are pessimistic or negative toward education, they are much more likely to say they believe there is not enough support in the community. This view was held by nearly 75% of caregivers in Ghana, and sizeable minorities in the other project country areas.

---

[65] Unlike the evidence for other barriers, these results are presented as a bar chart rather than line charts. This is because there is very little meaningful change from endline to baseline, and the more interesting aspect of the data is the total prevalence of certain views.
[66] Ghana 2015: 13.23%, Nigeria 2011: 8.30%, Kenya 2009 3.34%, source: http://data.worldbank.org/indicator/SE.TER.ENRR.FE?locations=KE-GH-NG

One possible way to interpret this evidence is that the primary caregivers themselves are not the ones who have negative views about girls' education. It could be that other members of the household have more negative views (e.g. it could be that the primary caregiver is usually the girl's mother, while the father is less supportive of girls going to school). Alternatively, it could be that the people surveyed simply are not willing to admit that they have reservations about girls going to school. In either case, this general picture suggests **that approaches to changing attitudes that are not targeted at particular individuals or focused on specific issues will not be successful.** Since the members of the communities believe at a minimum that it is important to say that girls' education is important, any meaningful effort to change attitudes would need to identify a more specific problem.

Figure 37: Caregiver Perceptions of community support for girls' education
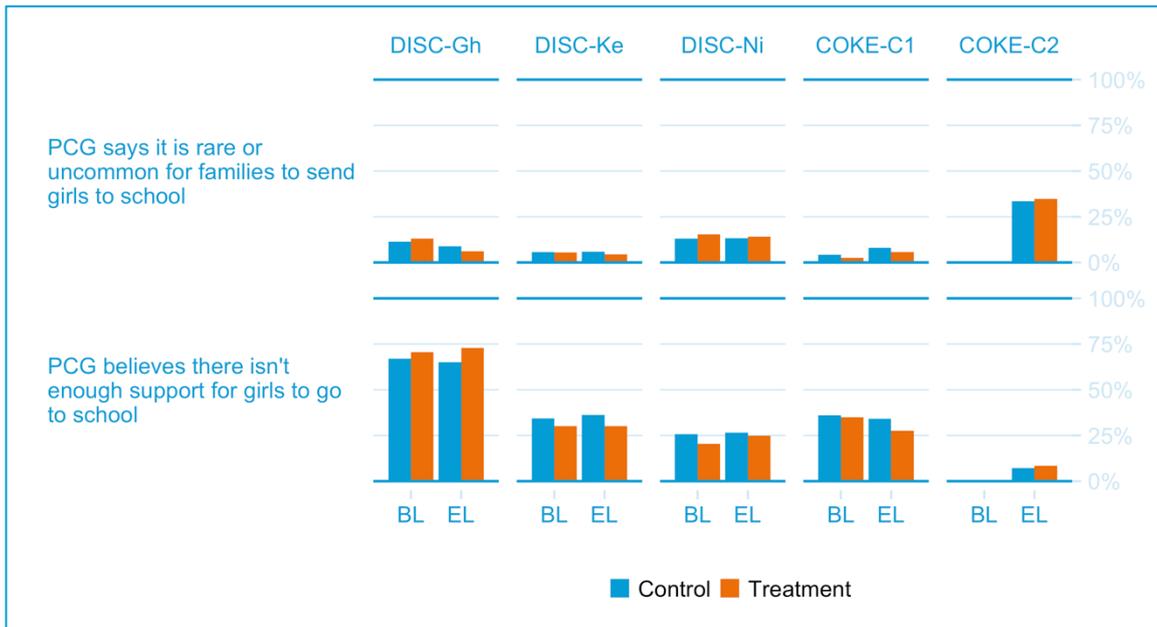
Considering some more specific indicators of attitudes, there is mixed evidence about how things have changed since baseline. One of the only areas where the project data shows a potential impact of the projects on attitudes is in parental involvement in education.

Figure 38 shows families where no member is part of a school committee. Participation rates vary significantly across project areas suggesting that its significance depends on the context. Moreover, while the students in Kenya had the best outcomes at a country level, very few families participated in school committees in the Kenya project. In this context, it is notable that Ghana started off with a dramatically higher participation rate than the other project areas, and the treatment group's participation improved more than the control group. Considering this alongside the evidence that families in Ghana had a strong belief there is not enough community support, there appears to be a significantly different social context in Ghana than in Kenya in particular: families appear to be highly engaged but frustrated with the support available.

Figure 38: Evidence of parental involvement in children's education
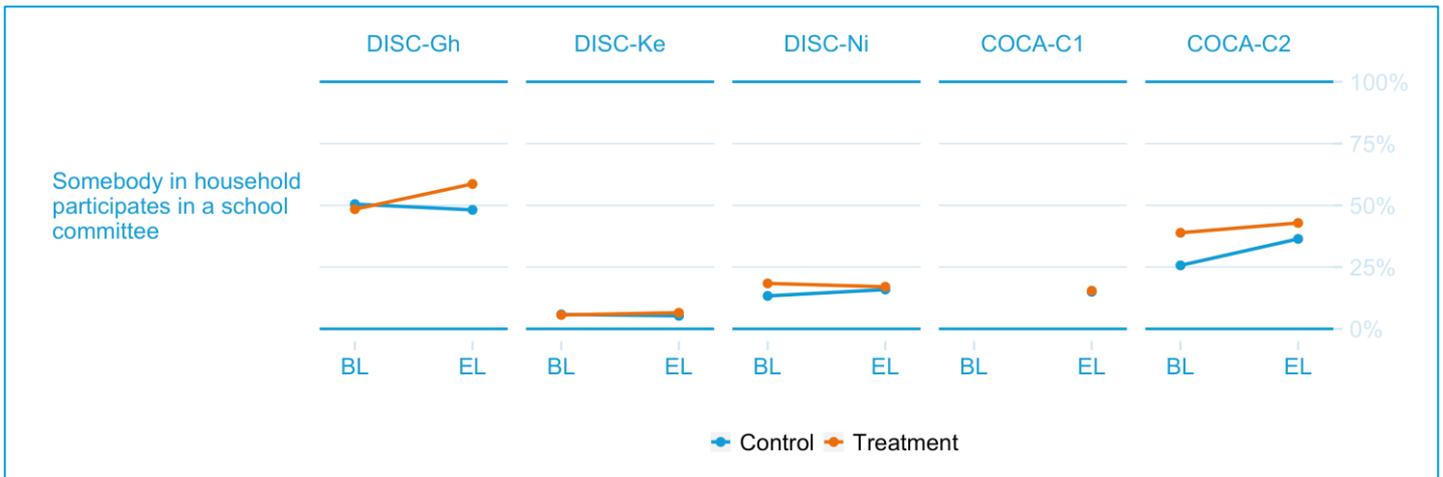
Finally, while we generally have little direct evidence of the effectiveness of particular interventions attempting to change attitudes, one exception is the television programme produced by the Discovery project. In their household survey, Discovery asked families if they had seen the programme they produced. The results are that almost no households had seen the programme. There are a number explanations for the lack of viewership (discussed previously in this section). In Figure 39, we examine one problem, which was that many households simply did not have a television. The most household-owned televisions were in Kenya and viewership was higher there, but even then less than 5% of households had seen the programme.

Figure 39: Reach of Discovery television programme and proportion of families who do not own a television
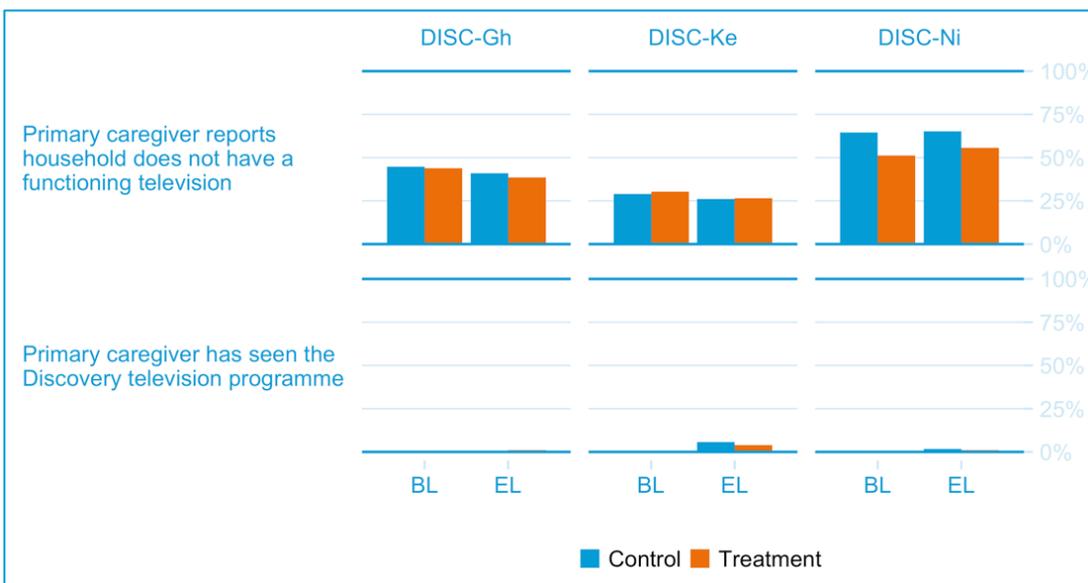
Have changes in barriers had an effect on education outcomes (attendance, learning)?

As shown in Annex E Table E.16, projects have not clearly demonstrated changes in attitude-related barriers to education. However, Discovery identified linkages between its interventions and educational outcomes and found evidence that in Nigeria and Ghana, parents' engagement in girls' education had a positive effect on their reading and maths scores. In Wajir, it presented qualitative evidence that parents' engagement possibly had an effect on the girls' attendance in school.

Similarly, Coca-Cola provided some qualitative evidence suggesting that that the engagement of community members possibly has a positive effect on girls' attendance in schools, particularly as these community groups would monitor the girls' attendance and follow up with her in cases of absenteeism.

3.4.5 Violence

> **Key Findings**
>
> At baseline, both Discovery and Avanti found evidence of violence-related barriers to education that exist in their targeted communities. Only Discovery aimed to indirectly address some of these factors through gender-responsive training for teachers and through community workshops. There is no evidence that these activities have been effective as Discovery provided evidence that violence, particularly in terms of corporal punishment and safety during the journey to school still existed, but did not investigate the change in this barrier over time. There is no evidence provided suggesting a change in education outcomes in relation to interventions addressing violence-related barriers.

What have projects found at baseline?

As shown in Annex E- Table E.17, only Discovery anticipated violence-related factors. However, during the baseline research, both Discovery and Avanti found violence-related barriers to education in their targeted communities. Coca-Cola did not make any assumptions about violence-related barriers or report on them at baseline.

Discovery anticipated sexual harassment and abuse to be a barrier to girls' education in all three countries, but found clear evidence of its existence only in Ghana. However, it did find some evidence that violence and harassment in Nigeria and Kenya existed, although it was not specific to sexual violence and harassment. In Nigeria, this was described in terms of corporal punishment by teachers. Avanti found that insecurities and fear related to the long commute to school was a barrier to girls' education.

Have interventions been designed to address the identified barriers?

As shown in Annex 3- Table E.18, Discovery did not design direct interventions to address student safety and security issues. However, it tried to address the topic through gender-responsive training for teachers and community workshops. There is no evidence that Coca-Cola and Avanti designed or implemented activities that aimed to specifically address violence-related barriers to education.

What has changed since baseline?

At endline, Discovery and Coca-Cola (Cycle 2) provided some evidence with regards to violence-related barriers to education. In all three of Discovery's countries some qualitative evidence was provided (in Kenya the evidence is clearer in Nairobi than Wajir) that **corporal punishment** still exists within the targeted communities and that some girls still feel **unsafe during their journey to school**. However, the project did not investigate changes in these barriers.

Indirect interventions to address violence-related factors have not been effective.

In all three countries, Discovery aimed at indirectly addressing violence particularly at school and by teachers through teacher training and community workshops. These however do not appear to be particularly effective because in all three countries there is evidence that corporal punishment by teachers still exists.

In Nairobi (where corporal punishment is unlawful[67]), the report found that 14% of treatment girls felt afraid at school, mainly due to teachers' behaviour and corporal punishment. In Nigeria (where corporal punishment is still lawful[68]), the ratio was significantly low with 5% of girls feeling insecure at school or on the way to school. In Ghana (where corporal punishment is lawful in schools[69]) about 4% of the girls felt afraid at school and particularly referred to corporal punishment by teachers.

Additionally, there is evidence that bullying, especially by boys, is another reason why girls feel insecure at school or on the way to school. This is despite the fact that project staff during interviews mentioned that tackling harassment by boys was also addressed by designing special toolkits for teachers and mentors. There is no further evidence to assess the effect of these efforts.

---

[67] endcorporalpunishment.org
[68] endcorporalpunishment.org
[69] endcorporalpunishment.org

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

Many external events independent from GEC activities occurred during the last two years of implementation that directly or indirectly affected project delivery. Each project intervenes in a given political context, meaning that events such as elections triggered short-term changes in government activities and/or more permanent shifts such as political reforms. SPW projects had to quickly adapt in order to avoid important delays in their implementation.

Similarly, the presence of other projects operating in GEC areas has had an impact on SPW projects' outcomes, either by reducing their relative impact in treatment areas when non-GEC projects improved the situation in control areas, or by helping GEC projects attain their objectives in treatment areas. Violence and environmental disasters were also significant obstacles to implementation, leading to temporary closures of schools or rendering some areas inaccessible.

The presence of other non-GEC projects in GEC areas contributed to improving girls' learning.

Other projects operating in GEC areas also influence GEC projects' work. These external interventions can positively or negatively affect SPW projects' results. For example, some treatment schools in Nairobi suggested that external support in the form of feeding programmes, provision of sanitary pads and bursaries might have contributed to the increased enrolment in their schools. Other projects were also reported to have been implemented in control schools, carrying out similar interventions to Discovery's. For example, in Nairobi, Plan International, Dignities, Save the Children and Population Council, and TUSOME were all reported to have been carrying out similar activities to Discovery. Specifically, they carried out teacher training (including in pedagogy, student-centred teaching and gender responsiveness pedagogy), cash transfer activities and the provision of laptops and interactive projectors. In Wajir, Save the Children was also reported to have been implementing a conditional cash transfer programme, which affected children's attendance in schools.

Avanti also reported that a number of projects were being implemented in their treatment schools at baseline. These activities were described in terms of feeding programmes and the provision of sanitary pads to girls. These were perceived as complementary to Avanti's project and as positively impacting Avanti's results.

Coca-Cola did not report on non-GEC similar interventions in its targeted communities. However, during interviews, it was mentioned that DFID funds other education projects in Nigeria and in the targeted communities, and that it facilitates meetings to coordinate between these projects. However, it was not clear to what extent these had an effect on Coca-Cola's interventions or results.

Academic year disruptions due to teachers' strike and environmental factors including drought and Ebola.

Discovery and Avanti reported that in Kenya there was a highly disruptive national teachers' strike for a whole month between September and October 2015. As described by Avanti, the strike led teachers to focus on finishing the school syllabi in time for the students to sit for their exams. This minimised the exposure of students to the learning content. For Discovery, this was also described in terms of parents pulling their children out of formal schools and enrolling them in non-formal schools to avoid prolonged school closures.

In Nigeria, Discovery also reported that the national elections in early 2015 and the Ebola outbreak in late 2014 led to complete school closures (during the Ebola outbreak this lasted for two months between September and October 2014), and Ramadan led to shortened school days.

Violence directly and indirectly affecting projects' implementation and results.

Schools across Nigeria were affected by the Boko Haram attacks. While this was not particularly explained in the endline reports, staff mentioned in interviews that these affected the attitudes of parents towards girls' education because they feared for their girls' safety and security. This generally makes work on parents' attitudes a challenge.

Discovery and Avanti also noted that some of their treatment schools were subject to robberies and that equipment was stolen from schools. According to the endline report of Avanti, this happened to one of its targeted schools in Kenya, and according to Discovery's completion report, it happened to five of their schools in Nairobi. Discovery's staff reported in interviews that in Wajir, security and safety issues were particularly evident. This resulted in trained teachers leaving Wajir (there are not many teachers from Wajir), so when the insecurity levels spiked, teachers left and did not want to go back to teaching at their designated schools.

Lessons learned

School barriers

- **Solutions provided by Discovery and Avanti did not address existing school barriers** such as the poor infrastructure (lack of desk and chairs in the classrooms, leaking roofs that could damage the equipment), and teachers' skills to use these technologies.
- **Both Discovery and Avanti provided evidence that teaching improved due to the use of ICT equipment**. ICT tools and content reduced teachers' workloads to a more manageable level. It also made teaching easier – particularly when explaining abstract ideas. However, it is necessary to ensure that teachers are able to use the technology and to integrate it into their teaching practices.

Poverty

- **Stipends appeared to cover a large part of school-related expenses**. Although these stipends also seemed to have a positive effect on attendance, this effect is unlikely to be sustained, particularly if the payment of stipends ends, which is what was reported.
- **Income-generating activities may have helped to reduce poverty but do not directly impact on learning outcomes**. Particularly with older girls, the additional source of income did not lead to an increase in expenditure on education-related costs because girls were keener to use this money to help their families rather than to pay for education.

Aspirations

- **Girls' clubs were an effective intervention enabling girls to enhance their confidence and identify their aspirations, but there is no clear relationship between an increase in self-confidence and improvement in educational outcomes**. Avanti and Discovery faced challenges in implementing these clubs because of a lack of support from parents and teachers, as well as a lack of financial support (especially for teachers providing additional support beyond their normal teaching duties). Furthermore, recruitment to these clubs was not random. It was often the most motivated girls who participated, performed the best and received the most support at home.

Negative attitudes

- **There is mixed evidence with regards to the effectiveness of the community outreach and awareness-raising activities in improving attitudes and support for girls' education.** Indeed, changing anchored social norms and perception of gender roles in the society is difficult to achieve, especially in such a short time period. Moreover, in a context of poverty, parents prioritise boys' education over girls' education. Parents' engagement was limited to those who already had an interest in education (often those who are qualified and educated to a certain level). On the other hand, it was difficult to engage with parents who lack household income and resources.

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

Key Findings

Although there is some evidence that communities are committed to continuing to support the delivery activities after the end of the SPW projects there is also strong evidence that **communities do not have the financial resources** to accomplish this objective.

Avanti and Discovery in particular **rely on schools** to provide the space and teachers needed to continue using various learning resources the projects have introduced and implemented. However, sustainability strategies that rely on teachers and schools are undermined by high teacher transfers and a lack of financial support from communities and ministries of education. All SPW projects have **engaged government bodies** at national, regional and district levels, but this engagement has **not translated into concrete plans to fund projects' activities** so that they can continue.

All three SPW projects will continue to receive support (in one form or another) through the successor programme to the GEC, the **GEC Transitions (GEC-T) programme**. Without this support it is highly unlikely that many SPW activities would have continued as they stand now. The GEC-T aims at continuing to support the same beneficiaries as in the GEC and, as Discovery and Avanti girls will transition from primary to secondary school, there is a risk that resources previously allocated to primary schools during GEC (2014-2017) will be partly or totally reallocated to secondary schools resulting in a lack of ability to sustain the projects' activities**.**

What sustainability outcomes were expected from the SPW projects?

The SPW was designed to attract private sector partners who, with the FM and DFID, could co-develop solutions to problems that marginalised girls faced in accessing and achieving a quality education. DFID's underlying rationale behind partnering with private sector actors is that these partnerships enable DFID to attract sustainable private sector investment to support social development (education) policy and programme objectives. The success of these partnerships depends on the solutions that are developed delivering both the GEC's education outcomes as well as delivering commercial benefits for the private sector partners – particularly to deliver activities that can be sustained beyond the life of the GEC. The GEC Business Case (2012) set out the following expectations that specifically relate to sustainability in the SPW:

- **The SPW would *leverage private sector involvement and resources.*** Strategic partners from the private sector would match-fund (in cash or in kind) 50% of DFID funding up to £15m to *explore new ways of supporting girls in education.* Private sector leverage was identified as one of the seven critical success criteria.

- **Match funding from private sector partners was considered an important strategy for addressing a *significant funding gap for basic education*** across the global community*.*

- **Public-private partnerships were considered as a particularly effective way of ensuring sustainability by targeting new investment in the non-state sector.**

- **However, sustainability was considered the most difficult aspect of the design of the GEC.** DFID envisaged that *by deciding to design the GEC to operate <u>beyond 2016</u>, although not allocating project funds, it is more likely that the impact of this funding can be sustained into the future as there will be ongoing management, and monitoring and evaluation of the GEC.* This suggests that sustainability would only be achieved if the programme was extended beyond the programme period.

- The Business Case recognises that for all projects (not just SPW projects) '*achieving sustainability (systemic change) from interventions is very hard to do, and so there is a risk that funds naturally flow to organisations who are able clearly demonstrate short-term results but are less able to articulate the long-term sustainability of interventions*' – this highlights the potential trade-off that projects may make. To mitigate against the risk that sustainability would be sacrificed for short-term results, the Business Case states that '*in practice, there will need to be an understanding of how funding will be sustained beyond the programme period, which will require some implicit commitment to sustain support until alternative sources are in place'*.

It's worth noting here that the whole GEC programme (including the SPW) finished in March 2017. **GEC Transitions (GEC-T)** has been developed and set up by DFID and the current FM as a successor to the GEC. It will provide further funding and support for those GEC projects that are able to meet and deliver GEC-T's objectives. As such, GEC-T potentially provides SPW projects with more funding (and time) to enable them to find alternative resources to continue and sustain the activities and benefits delivered during both the GEC and GEC-T programmes.

Box 11: Endline guidance for SPW projects on reporting sustainability

Before starting their endline research, the EM and FM provided SPW projects with guidelines for the endline analysis and reporting. These were not prescriptive, but they suggested that projects should report on similar areas that SCW and IW projects were required to report, as set out below.

What is the project's sustainability strategy?

- Provide a brief summary of the project's long-term sustainability and succession strategy including **commercial strategy, planned activities** and any progress to date in delivering it.

- Explain the **financial nature** of the partnership developed. Explain which activities were funded by the partners and which were funded by DFID, and **how the project will sustain itself with the ending of the matched funding** by DFID

To what extent has the project identified the pre-conditions for scaling up and /or sustaining its activities and results?

To answer these questions use analysis of quantitative and qualitative data and the project staff's own understanding of the project/context.

- How have contextual factors (e.g. environment, political, country context, education policy, financial resources, local government support and community) affected your capacity to deliver results that are sustainable? Consider:

  o The opportunities for sustainable change – be specific to your project.

  o The constraints and risks to sustainable change – be specific to your project.

  o Specific factors within the education system that are enabling or hindering sustainable change.

How has the project strategically engaged with other organisations to achieve complementary effects?

- Who has the project engaged with and to what effect?

- Has the project influenced changes in the behaviour and practice of others?

To what extent has the project leveraged additional investment?

- Summarise any match-funding, additional investment or in-kind support secured and/or received by the project.

What are your plans for delivering sustainable results?

- How does the project plan to overcome any challenges to sustainability to continue to deliver its outcomes after the end of the GEC?

What are the lessons learned about the scalability and sustainability of the activities delivered?

- Provide key lessons learned about the scalability and sustainability of the project's activities. The formulation of the lessons learned should follow the same approach as that set out in Section 4.5 and consider:

  o What you will or would do *differently* now or next time and why?

  o What you will or would do the *same* and why?

In 2016-2017, SPW projects developed their Sustainability Strategies and Plans following the Sustainability Sand Box event organised by the FM. We reviewed these plans for this endline evaluation. We have based our window-level findings on the basis of the evidence presented by SPW projects' external evaluators about the extent to which these plans had been delivered at the time of their external endline evaluations. None of the reports followed the guidance provided in Box 11. Instead, the external evaluators chose to evidence and analyse sustainability in different ways and to varying degrees.

As mentioned previously, out of all three SPW projects Discovery most extensively covered sustainability in their endline evaluation reports and provided the most useful critical insights into the opportunities and challenges that the project faces with regards to sustainability. We used these valuable insights to highlight typical characteristics of the sustainability of SPW projects, but this does not mean that Discovery's project was more or less sustainable than other projects.

Annex E.2 provides summary tables of the evidence about sustainability that we extracted from each of the project's endline evaluation reports. We used the following criteria to extract and analyse the projects' sustainability evidence:

- Information about which **type of activities proved most effective** and as a result should be sustained;

- The cash or in-kind **resources that the project has leveraged** specifically for the purpose of sustaining activities after the end of the project;

- Information about **who would be responsible** for sustaining key activities;

- Evidence of **key drivers** sustaining the project's activities;

- Evidence of **key barriers** to sustainability; and

- A **judgement** about how likely each of the activities identified would be sustained given the evidence available.

We also categorised different parts of projects' sustainability strategies at global, national, regional, school and community levels.

Key findings across the SPW

*All projects reported that they have **engaged and mobilised communities** to support the delivery of their activities and there is some evidence that communities are committed to continuing to support them after the end of the projects.*

Avanti reported that some parents had made contributions to cover the cost of electricity and to improve the ICT labs and school infrastructure. A majority of Discovery's schools across all three countries have set up Learning Centre Management Committees (LCMCs) who have developed Management and Sustainability Plans, which also involved training a community member to identify barriers to sustainability and develop a Community Action Plan (CAP). Like Avanti, Discovery also reported communities investing in the school infrastructure needed to sustain the project's activities:

> *"In leveraging resources, DP [Discovery Project] has had the most success in inspiring the community to secure and power the LC equipment. The community appears interested in continuing DP particularly since they directly see the benefits of the program. There was evidence that LC security has been considered a top priority for some communities and has been included in their Community Action Plans (CAP) and the community has been asked to contribute funds to secure the room with metal doors. Also, some schools through the school committees were able to raise money to pay the school's electricity bill, pay for a security guard, and to buy fuel for the generator."*

Both Avanti and Discovery adopted a whole school approach to delivery, whereas Coca-Cola specifically targeted its support to individual girls. As a result, Coca-Cola's endline evaluation reports mention the role of communities in sustaining their activities much less than the other two projects, but they reported that it was quite likely that the *improvement of the behaviour and attitudes of Gatekeepers and community members to girls' education would be sustained based on positive responses received from the qualitative analysis.* Coca-Cola's endline reports mentioned few if any barriers to sustainability, which we believe is a result of under-reporting.

*While there is evidence of community support and some commitment to continue projects' activities, there is also strong evidence that **communities do not have the financial resources** to maintain and continue current activities.*

In Avanti's and Coca-Cola's endline evaluation reports there is little or no explanation about how communities and in particular the relatively new structures that have been set up, such as LCMCs, will continue to develop and maintain the level of commitment and resources required to sustain time and resource-intensive activities. Discovery's endline evaluation reports provide the greatest insights into these challenges, citing issues such as: parents in Nigeria and Ghana not being able to continue paying to maintain or replace deteriorating equipment; diminishing enthusiasm and increasing tensions among communities; and crucially communities reporting a lack of clear plans in place setting out how communities should continue activities. Discovery Nigeria's endline report summarised this situation as follows:

> *"It (Fitila) enlightened our teachers and even our students now understand more than you can think of. In the past, there is the problem of absenteeism, but the coming of Fitila project has made the students to be punctual in attending school. Also, as a result of the activities brought by Fitila, instead of being told stories in the past, where you only use your ears, now, your ears will hear and your eyes will see. … if you say Fitila will stop, we will not be happy. This is because all the things that are done are done with funds. … we cannot do these things; the likes of buying a generator, buying CDs… we will not be able to afford buying the fuel."*

Coca-Cola specifically targeted individual girls rather than whole classes. Therefore, its sustainability strategy has a much greater focus on the **resilience of individual girls and groups of girls to continue benefiting from capacity gains** they have already made as a result of participating in project activities. For example, girls were provided with seed grants for business start-ups and to expand and buy equipment, and they were trained in micro-franchising /vocational skills. Similarly, the savings groups were provided with tools, helped to open bank accounts and to explore becoming registered associations. However, as with new community-based structures, these capacities are relatively immature and fragile. **It's not clear what ongoing support girls and groups can call on to help them become stronger and more resilient to shocks and setbacks** if the project no longer existed. New structures and capacities need time and support to become sufficiently strong to continue to develop and evolve.

*Avanti and Discovery in particular **rely on schools** to provide the space and teachers needed to continue using various learning resources the projects have introduced and implemented. However, sustainability strategies that rely on teachers and schools are undermined by high teacher transfers and a lack of financial support from communities and ministries of education.*

Avanti reported that teachers were needed to act as champions for using the projects' learning resources. Coca-Cola reported that 50% of their target schools had replicated their safe space models. Discovery reported across all of its countries that it had set up and trained resource teachers to lead and provide peer support to other teachers. However, Discovery and Avanti endline evaluations reported that **high teacher transfer rates** undermined any reliance on teachers to continue changes in teaching practices that had been implemented during the life of the project, as highlighted in Discovery Nigeria's endline report:

> *"Teacher transfers were often cited as a threat to the sustainability of DP activities, particularly when trained teachers are transferred to another school and not replaced with other trained teachers. A head teacher explained that: "Teachers are always being transferred. They are here today and they are in another place tomorrow. The ideal thing is the project should benefit all the students of all schools, not only schools where there is Fitila"*

Avanti reported that 83% of Head Teachers surveyed stated that they were willing to make contributions to costs to continue the project through: (a) parental contributions (80% of respondents); and (b) MoE contributions (29% of respondents). While this shows commitment from Head Teachers, it is undermined by evidence that communities cannot afford to continue contributing to project activities, but also that there was **little or no evidence reported of MoE plans to continue funding** these types of activities, as reported in Discovery Kenya's endline evaluation:

> *"Data from several key MoE staff interviewed by SI showed that, despite DP's engagement with MoE towards sustainability, the MoE could not provide any concrete response regarding their plans for continuing DP and they were not aware of any plan in the works. One MoE staff said: "I don't know, that is at policy level, now that is not mine. I cannot assume such roles."*

*All SPW projects have **engaged government bodies** at national, regional and district levels with some success in terms of aligning with policy and influencing policy change. However, this engagement has not translated into concrete plans to fund projects' activities so that they can continue. This takes a lot of time and advocacy.*

Coca-Cola's endline evaluation reported that it has worked with Central Bank of Nigeria (CBN) on the financial literacy curriculum who have now committed to rolling this out across 36 states – financial literacy has also been incorporated across all government schools in Kano. Coca-Cola also reported partnering with the National Youth Service Corps who have agreed to deploy volunteers to schools to support the ENGINE model for safe spaces in schools. Avanti's endline evaluation reported engaging with County MoE officials from the start and working closely with them as gatekeepers to ensure that the project aligned with MoE activities in schools e.g. delivering iMlango over DLP tablets. Discovery's endline evaluations also reported that it had built government capacity to sustain and integrate new teaching and learning practices, but its endline evaluation for its project in Ghana highlighted the challenge facing any strategy that is reliant on government funding, which is that **changing government budget allocations at any level is complex and difficult**:

> *"You see GES is one of the Ministries or departments that is taking the chunk of national budget and so what we are proposing that they should do, involves cost and so that will mean more cost to GES so that is going to be their first challenge. First the infrastructure like having the video facility in every class in every school because you don't do for some people and leave others that will not be equity… so if we are going to implement it we do it nationwide and that means cost and that means the cost is not only going to be in material provision, it is also going to be in training of the teachers. We have to give orientation to our teachers as to how to use those things… if we are going to do it at their school level, it involves cost let alone we have to bring them together at zones and train them. And so people will move from their homes, we have to feed them, we have to accommodate them, and we have to look for facilitators pay them and all those things. So the cost financial constraints will be one of the biggest challenges."*

At this stage there is no evidence that SPW projects have influenced government ministries to directly fund their activities so that they can continue beyond the life of the projects.

*All three SPW projects will continue to receive support (in one form or another) through the successor programme to the GEC, the **GEC Transitions (GEC-T) programme**. Without this support it is highly unlikely that many SPW activities would have continued as they stand now. The extent to which girls, schools and communities will continue to benefit from GEC-T will depend on the subsequent design of the SPW projects which are now tasked with continuing to support the same beneficiaries supported in the GEC – there is a risk that resources will be reallocated to support older age groups to ensure existing beneficiaries continue to benefit while activities previously supporting younger age groups in the GEC may not be sustained.*

Avanti and Discovery will receive further funding through GEC-T. Coca-Cola is not continuing, but their partner Mercy Corp will receive further funding through GEC-T to continue Coca-Cola's ENGINE programme. Like the GEC, GEC-T has been designed to deliver improved learning outcomes and sustainability. More than the GEC, GEC-T has a focus on improving girls' transition, both between grades and across school phases, including beyond secondary school and into further education or employment. GEC-T is also committed to continuing to support the same beneficiaries supported by the GEC to ensure that they are able to complete a full cycle of education.

Projects supported through the GEC-T will need to redesign their projects to some extent to ensure that they are able to continue supporting the cohort of beneficiaries from the GEC as they grow older and transition to the next stage of their education or employment. At this stage in the GEC-T we do not know what the SPW project designs are and so it is not possible to assess the extent to which GEC-T will enable:

a) the type and focus of the activities delivered by SPW projects in the GEC to continue;

b) the benefits realised by girls supported by SPW projects in the GEC to continue to benefit as they progress; and

c) the innovations used by SPW projects to be replicated and scaled up in other places.

As the original DFID Business Case explained, it was always a challenge to expect activities to be sustained over such a short timescale. The SPW has proved to be no exception. **Without continued support through GEC-T the projects did not report any evidence to suggest that their activities would continue beyond the short term.**

GEC-T is focused on supporting the same cohort of girls supported in the GEC as they progress their schooling or transition to higher /further education or employment. This suggests that SPW projects will need to ensure that their activities are directly meeting the changing needs and priorities of existing beneficiary girls as they get older. This raises a question about the extent to which girls who are not existing beneficiaries will continue to benefit from the type of activities that were provided in the GEC, but may not be supported in GEC-T.

There is a risk that children and girls in grades supported in the GEC will no longer be supported in GEC-T and new children entering these grades will only continue to benefit from GEC funded activities if they are sustained by the school /communities or through continued support from the GEC-T programme. If they do not continue to receive support through GEC-T then is it likely that projects' activities previously supported through the GEC will not be sustained. For example, Avanti and Discovery previously only supported girls and children in primary school. Assuming that project implementation on GEC-T will start in 2018, the GEC beneficiaries will range from Primary Grade 4 through to Junior Secondary Grade 10 in Kenya. If projects are only supporting previous GEC beneficiaries then Grades 1-3 would no longer continue to receive financial support through GEC-T. The learning resources provided for these grades in the GEC would need to be self-sustaining to ensure that children continued to benefit beyond the end of the programme. The extent to which this risk is likely to happen depends on how targeted SPW project designs are in the new GEC-T programme.

Lessons learned

- **Engaging and mobilising communities to support the delivery of projects' activities was related to a commitment to support the activities after the end of the projects**. However, there is strong evidence that communities targeted by SPW partners do not have the financial resources to ensure the sustainability of the projects. Avanti and Discovery in particular rely on schools to provide the space and teachers needed to continue using various learning resources the projects have introduced and implemented. Sustainability strategies that rely on teachers and schools are also undermined by high teacher transfers and a lack of financial support from communities and ministries of education.

3.6 Value for money of the SPW

Investment in the SPW

Table 21 below provides an overview of the total investment made by DFID and each of the partners in the SPW projects covered by this evaluation (and Ericsson's project in Myanmar, which is not covered by this evaluation).

Table 21: Total investment in each of the SPW projects

SPW project	DFID funding (£) [1]	Partner funding (£) [2]	Total (£) [3]
Discovery Communications	£13,032,426	£12,304,722	£25,337,148
Coca-Cola	£3,964,227	£3,499,292	£7,463,519
Avanti Communications	£14,141,627	£13,765,663	£27,907,290
Sub-total:	£31,138,280	£29,569,677	**£60,707,957**
Ericsson	£3,731,409	£3,659,766	£7,391,175
Grand Total:	£34,869,689	£33,229,443	£68,099,132

Notes:

[1] Actual total DFID expenditure for whole life of the projects as reported by the FM on 06/07/2017.

[2] Partner funding figures sourced from SPW summary in FM (2015) Strategic Partnerships Consultation Report – pending final FM match funding figures.

[3] Subject to change depending on final match funding figures from FM.

Table 22 below provides a breakdown of the costs by type of inputs across all three projects. Project outputs cannot be used to assess value for money because the way they are typically defined makes it too difficult to identify the type of activities or activities associated with each output. For this reason, the FM asks projects each year to submit value for money (VfM) metrics tables that categorise project costs specifically by type of input. Table 22 has been derived from the VfM metrics tables submitted by each of the projects at endline. Most of Avanti and

Discovery's expenditure relates directly or indirectly to the ICT focus of their projects, while most of Coca-Cola's expenditure relates to their focus on providing girls in school and out of school with courses to help improve their livelihoods and generate income.

Table 22: Breakdown of project expenditure of DFID funding by type of input

Project Inputs	Avanti		Discovery		Coca-Cola	
	DFID contribution	% of contribution	DFID contribution	% of contribution	DFID contribution	% of contribution
ICT infrastructure and connectivity Content development and licenses (including girls' clubs content) Maintenance Teacher Training	£10,361,145	73%	£9,695,138	74%		
Stipends	£1,608,967	11%				
sQuid cards	£884,839	6%				
Sustainability	£1,286,676	9%				
Community engagement, Girls' clubs			£1,719,679	13%	£488,860	12%
Mass media show (production and dissemination)			£1,617,609	12%		
Interventions with in-school girls (Safe Spaces)					£1,889,147	48%
Interventions with out-of-school girls (Safe Spaces)					£1,586,220	40%
Total DFID funding [1]:	**£14,141,627**	**100%**	**£13,032,426**	**100%**	**£3,964,227**	**100%**

Notes:
- Data source: End of project VfM metrics tables submitted to FM by each of the projects (June 2017).
- Breakdown of partner match funding is not available.
- [1] Actual total DFID expenditure for whole life of the projects as reported by the FM on 06/07/2017.

Cost-effectiveness of SPW's learning outcomes

The value generated by the SPW projects at the end of the programme is defined first and foremost by their impact on the GEC's education outcomes i.e. girls' literacy and numeracy outcomes. There is little evidence to suggest that SPW projects had a statistically significant effect on the literacy and numeracy outcomes of the girls they targeted and reached.

The only conclusive evidence of a positive impact across the SPW was in Discovery's literacy results in Nigeria with their target girls improving their UWEZO scores by 1.6 out of 40 above the control group since baseline, representing an achievement against target of only 71% (0.14 SD). However, by achieving 41% of its numeracy target, Discovery did not demonstrate any impact in this geographical area. As per the FM's beneficiary methodology[70] this level of achievement against the target translates into an average achievement of 56% which

---

[70] Extract from PwC Annual Report Methodology Note Appendix: *The FM reports learning results that have been calculated using project data. Learning results are converted to the number of learning beneficiaries by scaling their performance against their targets and scaling this to the overall beneficiary number. This means that both projects which have met the set target and projects which have not met the set target but out-performed the control group are able to report benefits in terms of girls' learning gains. The number of girls who learned is linked to the learning*

means that 88,556 girls (out of the 158,135 beneficiaries in Nigeria) benefited from improved learning. As noted below, the size of the target group is a significant determinant of the number benefiting and Discovery's approach involved a 'whole school' approach, which resulted in large target figures and a relatively large number of beneficiaries when scaled up[71]. There are no target figures for boys specifically. Discovery took a whole school approach to targeting and delivering its interventions so it is reasonable to assume that boys may also have benefited from the project's interventions, but it is not possible to estimate how many accurately.

The purpose of cost-effectiveness analysis (in this context) is to compare the efficiency of the ways in which different projects have used their resources to impact on the GEC's literacy and numeracy outcomes for marginalised girls. While the unit cost per girl benefiting from improved literacy, numeracy or overall learning could be calculated, there are no benchmarks available to compare Discovery's cost-effectiveness against which could provide meaningful findings or learning.

Cost-effectiveness of SPW's attendance outcomes

Across the SPW, there were issues with the quality of the attendance data and measurement of attendance. These issues limited our ability to generate conclusive findings about the impact of the SPW on girls' attendance at school. Avanti's attendance results were inconclusive. Coca-Cola provided attendance data for their Safe Spaces rather than the schools they supported. Discovery achieved its attendance targets at endline in Nairobi (5.1% above the control group – a 255% achievement of the 2% point increase target) and in Nigeria (2.2% above the control group – a 110% achievement of the 2% target). Coca-Cola nearly achieved it for their Safe Spaces for both cycles of the support they provided, but this relates to girls' attendance on the courses that had a relatively small focus on literacy and numeracy and as such cannot be compared with attendance at school, which is a key GEC educational outcome. There is no methodology for estimating the number of girls who benefited from improvements in attendance.

Reach and equity

SPW projects aimed to reach marginalised girls. According to the projects' endline reports, SPW projects **reached 380,350 girls and 426,920 boys.** Discovery considered all of the girls in their project areas to be marginalised and delivered activities based on a **whole school approach**. Avanti did mostly use a whole school approach to implementation but did not consider all girls to be marginalised, and the stipend component was used to target the most marginalised girls. To some extent this explains the large number of children reached and in particular the large number of boys participating in SPW activities. Avanti estimated that 10,497 families and 115 merchants have been reached whereas Coca-Cola estimated that they reached about 40,000 community members. There is limited information available from these projects at endline about whether they reached the most marginalised girls in their regions and communities, or which groups of girls make up the most marginalised in the project.

Coca-Cola was the only project that targeted specific girls for all of its project activities. Girls undertook self-screening to participate in the project based on criteria that included factors such as: being an orphan; a parent being sick; having a disability; or being married before the age of 18. The EM reanalysis of the data suggests that Coca-Cola had some success in reaching the most marginalised, particularly orphans /girls in single-headed households.

Value for money conclusion

The overarching conclusion (discussed in detail in Section 4) is that the SPW has generally not achieved its literacy and numeracy targets at the end of the programme largely because of weaknesses in projects' design and in the ways in which they were implemented. For example, while Coca-Cola's project improved financial literacy skills

---

target, usually 0.2 standard deviations. If a project achieves its target of 0.2 standard deviations, it would be assumed that all of the target beneficiaries have learned, and therefore the number of learning beneficiaries would equal the total number of target beneficiaries calculated as explained above. For example, for a project with a target of 5000 learning beneficiaries, meeting 100% of the target, the calculated value of learning beneficiaries would be 5000. If the project reached less than the target, the numbers would be scaled downwards in a linear fashion so if a project met half the target, then the number of learning beneficiaries would be assumed to be half of the total number of target beneficiaries. For example, for a project with a target of 5000 learning beneficiaries, which achieved 50% of their learning target (i.e. a 0.1 standard deviation improvement against a control group), then the calculated value for learning beneficiaries would be 2500. Both the literacy and numeracy results were combined and averaged so for example a project that achieved 100% of their literacy target, and 50% of their numeracy target, would have an overall learning outcome of 75% against target. So in the above example, for a project with a target of 5000 learning beneficiaries, the calculated value for learning beneficiaries would be 3750.

[71] The number of beneficiaries of Discovery's project in Nigeria represents 42% of the total number of SPW beneficiaries. This explains why the number of girls that benefitted from improved learning is so high (23% of the total number of SPW beneficiaries).

among its target in-school girls, the amount of time spent in each cycle helping these girls to improve their literacy and numeracy was far too little to have a significant impact. Discovery's evaluation reported that showing a video to up to 100 students per class, in small classrooms with no tables and chairs in some instances limited the effects on children's literacy and numeracy, especially if some teachers reported being reluctant to show some of the videos because they did not believe they were culturally appropriate.

At the end of the programme there was little reported evidence to suggest that concrete plans were in place to enable activities to be continued without further support from the SPW projects. All three projects will continue to receive support through DFID's GEC Transitions programme, which is the successor to the GEC. Without further support from DFID it is unlikely that these activities delivered through the GEC will continue. The recommendations that follow at the end of this report suggest ways of generally improving project designs and implementation to increase effectiveness and impact. Unless SPW project designs and implementation processes are better adapted to suit the specific learning needs and priorities of their target groups it is difficult to see how they will improve their impact and deliver better value for money education outcomes.

4 Conclusions

SPW projects had little or no impact on literacy and numeracy outcomes of the marginalised girls that they reached, largely because of a combination of weaknesses in projects' designs and the ways in which they were implemented, as well as contextual factors, such as the Ebola outbreak in Nigeria that affected implementation.

The evaluation of each SPW project's impact on learning outcomes was designed as a quasi-experiment. The main finding from this evaluation approach was that the literacy and numeracy results of project beneficiaries did not improve more than the results of comparable girls who were not beneficiaries. Across the SPW projects – disaggregating Discovery's project into the three project areas where it worked – the best performance was found in Nigeria with an achievement of 71% in literacy (statistically significant). By contrast, the worst performances found for numeracy were in Nairobi (Kenya) and Ghana where negative and statistically significant impacts were observed. Avanti had no impact on literacy and numeracy. Coca-Cola did not have any positive impact on learning, except for literacy in Cycle 1 where half of the target has been achieved (52%). Contextual factors, such as the Ebola outbreak in Nigeria may have affected both Discovery's and Coca-Cola's results over this period. It is worth trying to analyse the causes of the small positive impact in Nigeria's reading scores and the absence of results in the other geographical areas. In particular, the girls in Discovery Nigeria had by far the lowest reading levels at baseline, so one plausible interpretation could be that Discovery's approach can help with literacy, but only for students with the greatest needs. This resonates with the EM's findings from the midline evaluations of the SCW and IW that showed these portfolios as a whole had the greatest impact on girls with the lowest levels of literacy.

Taking the learning results for the window as a whole, we can report ten findings of no impact, two negative impact findings and one positive impact finding. This is broadly consistent with the hypothesis that SPW projects as a whole had no positive or negative effects on learning. Although there are issues with the quality of data, analysis and reporting, the lack of results reported by these SPW projects suggest that significant changes are needed to improve their impact on the marginalised girls they targeted.

Projects focussed on marginalised regions, not always on the most marginalised girls, but the evidence suggests that where interventions targeted more marginalised girls, activities have been more effective.

Projects applied different targeting approaches. While Discovery and Avanti targeted whole schools, Coca-Cola targeted specific communities. In the cases of Discovery and Avanti, the schools were identified based on their presence in marginalised communities. However, we found evidence that Avanti's targeted communities are no poorer than the national average and in the case of Discovery, some were also not particularly educationally marginalised. Avanti made a distinction between "educational marginalisation" and "poverty marginalisation" in their M&E plans using poverty as a primary factor to identify marginalised areas in Kenya. Furthermore, in order to ensure high level of marginalisation in the project schools, Avanti's sampling frame of schools included those that had access to electricity and participated in Kenya's sanitary towel programme[72]. However, restricting selection to those schools with access to electricity may have excluded some of the poorer areas from the sampling frame.

In cases where communities were educationally and economically marginalised, we found that some improvements in education outcomes have been achieved. Additionally, in the cases where interventions were more targeted, such as Coca-Cola's project, or the stipends component in Avanti's project, there is evidence that marginalised girls have been reached, and/or that some results have also been achieved as a consequence.

At the window level, the SPW projects have reached more boys than girls. However, there is no evidence that Discovery's and Avanti's projects led to an increase in gender disparities as a result.

Across the SPW, projects reached more boys (426,920) than girls (380,350). The underlying assumption behind the GEC is that girls face gender-specific challenges, which puts them at a disadvantage educationally compared to boys. Although the GEC Business Case states that the programme is designed to address disparities in girls' education compared to boys, it was not set up at the start to explicitly measure how or to what extent projects affected gender disparities in girls' education outcomes. The original Theory of Change for the GEC in the Business Case defined how the programme would be delivered rather than the changes that would be experienced by girls as a result. In this context, no gender studies have been conducted by any GEC projects at endline,

---

[72] Avanti's M&E framework.

including the SPW projects, and so cannot provide any conclusions about changes in gender disparities in education outcomes. Discovery's and Avanti's core activities predominantly involved a whole school approach using education technology. The ways in which these activities were delivered meant boys could not be excluded in the process of delivering support to the target girls. In this sense boys were reached and supported at no additional cost.

Projects did not focus the design of their interventions to achieve the educational outcomes targeted by the GEC.

Discovery's and Avanti's main project activities involved the introduction of the ICT equipment and improving pedagogy in schools. Coca-Cola provided business and finance training to out-of-school girls and additional instruction to in-school girls after school. These approaches responded to the expectation laid out in the GEC Business Case that private sector investment would *explore new ways of supporting girls in education*. However, the evidence suggests that these approaches were generally not sufficiently focused on improving school attendance and learning outcomes (i.e. literacy and numeracy) as intended by the GEC.

The largest group of beneficiaries in Coca-Cola's intervention was out-of-school girls. These girls did not receive any significant literacy or numeracy training. Instead, this intervention was focused on improving their financial literacy with further opportunities for potential employment in Coca-Cola's value chain, which could have created a disincentive for the youngest girls to return to school. Girls who were enrolled in school and were beneficiaries of Coca-Cola's intervention only received two hours of tutoring per week to cover all of the course content included in Coca-Cola's curriculum – this included business skills and finance, leaving less than one hour to cover academic content (school curriculum) including numeracy and literacy. From a programme design standpoint, this approach placed too little emphasis on the GEC's key learning outcomes.

The core elements of Discovery's programme focused on improving pedagogy – introducing ICT equipment to create a more engaging classroom, and training teachers in student-centred and gender-responsive teaching methods. However, neither the content introduced through ICT nor the new teaching methods focused on methods to improve specific skills in literacy or numeracy. Discovery's learning content, conveyed through educational videos, mostly focused on science and social studies rather than on literacy and numeracy. Teaching methods focused on student engagement and fostering creativity. While it is possible that improving the overall quality of teaching could improve specific outcomes such as literacy and numeracy, the project design, content or teaching methods did not focus on these outcomes. The lack of statistically significant impacts on literacy and numeracy at endline supports our conclusion that pedagogical interventions that are not specifically designed to improve literacy or numeracy are unlikely to have that effect. Avanti was the only project in the SPW that focused specifically on improving numeracy in their project design. Since we also found no impact of this project on learning outcomes, this is most likely explained by other factors relating to its implementation and context.

Projects did not sufficiently adapt their projects to the local context with respect to their design or implementation strategies.

Projects generally appear to have rolled out a blanket approach to designing interventions in different locations and contexts. There were some exceptions – such as Discovery's flexible approach to running the girls' clubs, or projects' community mobilisation and parents' engagement in education activities, which were generally driven from the bottom-up. The core activities designed to address education outcomes tended to be solution-driven involving particular types of equipment, products and solutions that projects had already developed and used elsewhere. The SPW challenged private sector partners to respond to the GEC with pre-conceived ideas about potential commercial solutions to the problems that marginalised girls faced. However, the evidence suggests that these solutions were not sufficiently tailored and adapted to suit the context, needs and priorities of the girls, schools and communities that projects were working with to have a significant effect on literacy and numeracy specifically.

For example, it is not clear how Discovery considered differences in contexts when rolling out its activities in its three target countries. For example, Ghana appears to be a special case when it comes to parents' engagement and perceptions about girls' education in comparison to the other countries and contexts where Discovery worked. The educational levels of schools were dramatically different across countries. At baseline, a large majority of students in schools in Nigeria could not read a single word, whereas students in Kenya were mostly able to read, but were somewhat behind international benchmarks. Moreover, the cultural context of the different areas varied greatly. Given these differences, it would be surprising if the same package of pedagogical techniques and learning materials would respond to the needs of the students in each area.

It also appears that the project did not take into account some of the target population's characteristics when the intervention has been designed. Discovery's National Talk Shows aimed to affect the attitudes of targeted communities around their perceptions of education. However, it appears that in rolling out these shows, basic considerations were not taken into account to ensure that the target communities could and would actually watch the programmes. The shows – in English, except in the Northern Nigeria where they were adapted in the main local language, Hausa - were often not in a language communities could understand, and the timing of the shows competed with popular soap operas. It seems that these problems were not anticipated in the design phase or addressed during implementation.

Problems in implementation of projects may also have reduced their effectiveness.

The evidence shows numerous challenges that projects faced during implementation. There is evidence that many of these persisted until the end of the projects. An example of this is Avanti's approach involved substituting the teacher with the individualised learning platform. This approach could potentially have delivered high quality content to students, however the project's evidence suggests that the approach did not get sufficient buy-in and engagement from teachers, so rather than substituting for teachers, the technology was simply underused by students.

There is also evidence that the number of hours per day in which the ICT equipment was functional decreased to an average of 2.6 hours per day (for a school of about 1000 student with only 25 computers). Some of the underlying causes for this included electricity bills increasing and schools not being able to pay them, or technical problems. The ICT equipment Discovery installed faced similar problems. Students were not always able to see the televisions, and the content was not always appropriate for what the teacher wanted to accomplish.

One clear finding from the SPW is that ICT equipment can only act as a tool to support teaching and cannot effectively substitute for good teaching practice. Ensuring that ICT equipment is effective appears to be a more intensive task than the projects anticipated, requiring monitoring and responding to the actual use of the equipment to ensure that it serves its intended purpose.

Projects' activities are unlikely to continue without further DFID support through the successor GEC-Transitions programme, and even then it depends on the design and reach of the new projects.

Sustainability was always going to be a challenge, but more than the other GEC windows, the rationale for the SPW was driven by the need to attract commercially sustainable private sector investment in girls' education. The SPW succeeded in attracting private sector investment but this investment has not delivered a sufficient impact on educational outcomes let alone sustainable activities or benefits. All of the projects effectively engaged government bodies at multiple administrative levels and they engaged, mobilised and sensitised schools and communities to the designs that they wanted to implement. At endline though, there are no clear concrete plans among communities, schools or government ministries to sustain continued investment in the activities specifically delivered by the projects. The projects' evaluation reports suggest that sustainability strategies are overly reliant on additional school and community resources that are fragile at best.

It is plausible that more time is needed as anticipated in DFID's Business Case, but arguably only activities that are effective should be sustained. Only those activities that deliver an evident benefit are likely to receive the financial support and investment in time and effort needed from stakeholders for them to continue. For example, Avanti's endline evaluation found that 34% teachers report using computers less than once /month while only 13% report using the internet to prepare for lessons. At the same time, low levels of time on task due to the lack of equipment per child was viewed as a key constraint to its effectiveness. Such a top-down approach to imposing poorly adapted solutions is unlikely to earn continued commitment from government stakeholders once the projects end.

All three SPW projects covered by this evaluation will continue to receive support (in one form or another) through the successor programme to the GEC, the GEC Transitions (GEC-T) programme. Without this support it is highly unlikely that many SPW activities would have continued. The extent to which girls, schools and communities will continue to benefit from GEC-T will depend on the subsequent design of the SPW projects which are now tasked with continuing to support the same beneficiaries supported in the GEC. There is a risk that resources will be reallocated to support older age groups to ensure existing beneficiaries continue to benefit, while previous GEC-funded activities supporting younger age groups or lower grades may not be sustained as a result of this change in focus.

It is plausible that with more time these projects could become more effective. But the evidence strongly suggests that significant changes need to be made to their designs and delivery processes and that more time in itself would

not make much difference. Only activities that are effective should continue and be sustained. In this respect, GEC-T provides the SPW projects with an opportunity to re-focus their designs and correct their delivery processes to have a greater impact on the marginalised girls they are targeting.

It is important to recognise that girls, their families and wider communities frequently bought into the projects, and the consequences of not sustaining activities that do deliver tangible benefits are significant for them, as expressed in Discovery Kenya's endline report:

> *"You see once you produce a child, you don't leave this child immediately this child learns how to walk. Help this child to mature more and more and then you say now you are above 18, now begin preparing on how you will move and settle as a grown up. So, I address them about this issue that I have made an appeal to them. Because it is about one year and something that they have been with us. If they will give us grace period to extend for a little bit more, yah until we properly stabilize then we'll say OK thank you."*

There were significant issues with the standard of quantitative and qualitative data, analysis and reporting across the SPW. A more prescriptive approach is needed to improve consistency and comparability.

External evaluators did not consistently use the household survey template provided by the EM because it was not mandatory to do so. As a result, the type and quality of data collected and submitted by each project varied greatly. Learning outcome test data was not linked to the household survey and demographic data that was collected. There was a lack of consistency in the questions asked by the evaluators, and key variables (e.g. attendance) were not collected consistently. Qualitative analysis in the evaluation reports was poor, providing limited explanations about how and why different types of interventions worked, for whom, in what contexts and with what effects. Despite a large investment of time and effort by the FM in particular, the learning from the evaluations is limited, lacks external validity and has made it very difficult for the EM to draw comparisons across the SPW and make generalisable observations about what has worked across the window. The EM's attempt to mitigate for this by interviewing key project staff did not provide sufficient information to fill the gaps in the evidence base or address the quality and consistency in the evidence submitted at endline. If external evaluators are to be conducting project evaluations in GEC-T then it seems clear that a more prescriptive approach is needed with greater technical oversight to deliver improvements in data quality, analysis, reporting, learning, consistency and comparability across the programme.

In the relatively short time that the SPW has been delivering its activities it has had little evident impact on the education outcomes of the marginalised girls that projects targeted. From a 'project' perspective, the designs and delivery processes need to change to have a greater impact, specifically on girls staying in school and improving their literacy and numeracy. In the context of developing new and innovative public and private sector 'partnerships' that focus on girls' education outcomes, DFID, the FM and the partners have progressed and completed the first phase of the GEC. However, the evidence from this endline evaluation suggests that these partnerships need to be rebalanced to place greater weight on the education outcomes of the marginalised girls that DFID is continuing to support through GEC-T.

5 Recommendations

The GEC programme came to an end in March 2017. DFID and the FM are currently in the process of developing project designs and planning baseline research for the next phase of the GEC. All three of the SPW projects covered in this report will continue to receive support through the new GEC Transitions (GEC-T) Window, which is one of two funding windows[73] that will make up the next phase of the GEC.

The following recommendations have been developed with a view to specifically informing the development of the SPW projects as they progress through the next phase of the GEC. These recommendations are aimed at DFID. However, the actions required to deliver them involve DFID, the FM, and the projects themselves. They are structured by specific recommendations for SPW projects followed by more general observations.

It is important to note that these recommendations, as with the evaluation report as whole, are focused on projects' effectiveness and impacts on improving girls' education outcomes. Where possible, the recommendations take into account the objective of the SPW to engage the private sector in new ways to deliver girls' education, but do not specifically relate to this because it is better covered by the findings from EM's Process Review[74] and the FM's Strategic Partnerships Consultation.[75]

1. **Coca-Cola and Discovery require a stronger focus on improving girls' literacy and numeracy outcomes** through teacher training, teaching content, use of ICT resources, and direct instruction. It appears that general improvements in the learning environment in Discovery's project areas has not translated into improved performance in the learning outcomes of girls as targeted by the GEC. Coca-Cola's direct instruction to students did not allow enough time to focus specifically on literacy and numeracy to improve these outcomes. Furthermore, out-of-school girls received no direct instruction to improve their literacy or numeracy.

   If the GEC-T programme's core objective is to significantly improve literacy and numeracy by the end of its lifetime then this should be the starting point for further problem diagnosis and key design decisions undertaken by projects. Project designs should be results-driven to ensure activities are relevant to the barriers to learning faced by the projects' target girls.

2. **Avanti and Discovery need to adapt their interventions to target girls more specifically.** Special tutoring classes aimed at improving girls' literacy and numeracy skills, as well as developing their aspirations for future, would be beneficial. These additional classes outside school hours could use ICT tools already provided by the programme. However, such interventions should be accompanied by extra financial support to the tutors or teachers for these extra-hours of teaching. Furthermore, the project should ensure that such classes are not clashing with girls' housework responsibilities by rolling out interventions aiming at reducing these duties. Avanti should also reconsider the provision of stipends to boys in order to reallocate these resources to activities targeting girls.

3. **Discovery needs to adapt its teacher training and use of ICT materials to better fit the local context.** Comparing learning outcomes across the three countries in Discovery's programmes suggests that the educational context is dramatically different. With extremely poor baseline outcomes in the Nigerian project areas, a focus on remedial education for basic skills may be necessary to continue to see observable improvements in basic literacy and numeracy skills. On the other hand, in countries where education systems are relatively well-developed, for example for in-service teacher training, parallel approaches to teacher training can be counter-productive for teachers who may become confused and resentful. In these contexts, more emphasis should be put on enhancing, complementing and supporting existing systems for teacher training, teaching curricula and existing ICT initiatives, in particular those being implemented by government.

   A more systematic approach to context analysis at community, school and institutional levels will help SPW projects ensure that their solutions and services are adapted to the needs and priorities of different target beneficiary groups, and partners and stakeholders whose participation is critical to their success. This has

---

[73] The other window is called 'Leave No Girl Behind', and focuses on those girls who are the most marginalised.
[74] Coffey (2015), GEC Process Review Report
[75] PwC (2015) Strategic Partnerships Consultation

to include sometimes very practical constraints related to the availability of electricity supply, internet connection, adequate classroom facilities and issues around affordability and maintenance.

4. **Discovery and Coca-Cola need to review the value added of their current initiatives to improve community support.** While the projects' evaluations found that many families feel there is not enough support for education in their community, they also provided little evidence that broad sensitisation campaigns had an observable impact on increasing community actions to support girls' education. Similar efforts to improve community support should be scaled back unless they can be designed to specifically deliver achievable and verifiable changes in community behaviour.

   Projects need a nuanced understanding of the complex relationships between different levels of community awareness of the benefits of girls' education and the practical decisions households take, to inform community-based interventions that are more specific and impactful as a result.

5. **All projects need to increase their support for poor families** to address poverty-related barriers or acknowledge that they are not able to support girls from poorer families within their target groups. All projects identified poverty as a key barrier to education. Relevant interventions that were delivered, such as providing stipends or supporting income-generating activities did not appear to be sufficient to tackle poverty-related barriers. Projects may choose to drop activities that seek to address poverty-related barriers and focus their efforts on improving the quality of teaching in school. However, if addressing poverty remains an important part of projects' intervention strategies then they should consider more direct and targeted interventions, such as cash transfers.

6. **If external evaluators continue to conduct project evaluations then a more prescriptive approach and greater technical oversight is needed to guide projects' commissioning of evaluations.** The quality of research and evaluation delivered by external evaluators needs to improve. DFID and the FM should be more involved in the selection of external evaluators. This should ensure that external evaluators are proposing the most appropriate counterfactual evaluation designs and mixed methods research needed to meet the evaluation and reporting requirements of GEC-T. Greater oversight by the FM of the commissioning process overseen by projects should also ensure that external evaluators have the type and quality of evaluation and research skills required.

   The Terms of Reference for projects' external evaluators should be more prescriptive, including requirements to: use a standard household survey template; link household survey data with learning outcome data; and collect specific disaggregated data about key subgroups. Furthermore, the FM could consider requiring projects to provide qualitative data, for example in the form of redacted transcripts (or notes) of interviews and focus group discussions. The intention is to improve the external validity of the quantitative and qualitative analysis and findings produced by external evaluators, and enhance the quality of learning from significant investments in these evaluations.

   DFID could also consider complementing the evaluations commissioned by the projects with primary research (largely qualitative research) conducted by an independent Evaluation Manager as a means of verifying in the field the reported changes, in particular how and why these changes have occurred.

   The overall aim is to improve the quality of data, analysis and reporting by projects to achieve greater learning, consistency and comparability across the portfolio.