
GEC-T girls' baseline learning levels and predictors of learning

Rachel Outhred, Maria Jose Ogando Portela, Alasdair Mackintosh, and
Katharina Keck

December 2020



About Oxford Policy Management

Oxford Policy Management is committed to helping low- and middle-income countries achieve growth and reduce poverty and disadvantage through public policy reform.

We seek to bring about lasting positive change using analytical and practical policy expertise. Through our global network of offices, we work in partnership with national decision makers to research, design, implement, and evaluate impactful public policy.

We work in all areas of social and economic policy and governance, including health, finance, education, climate change, and public sector management. We draw on our local and international sector experts to provide the very best evidence-based support.

About Oxford MeasurEd

Oxford MeasurEd is an innovative, global-facing consultancy that works with international partners to collect and analyse data to better understand and improve learning in low- and middle-income countries. We forge advanced new insights that shape the basis of education policies that facilitate lasting change and opportunity in the real world.

We understand that our clients often need information quickly, so we work to source data from a variety of sources, apply innovative data analysis techniques, and report on the strengths and limitations of the analysis.

Oxford Policy Management Limited
Registered in England: 3122495

Level 3, Clarendon House
52 Cornmarket Street
Oxford, OX1 3HJ
United Kingdom

Tel: +44 (0) 1865 207 300
Fax: +44 (0) 1865 207 301
Email: admin@opml.co.uk
Website: www.opml.co.uk
Twitter: [@OPMglobal](https://twitter.com/OPMglobal)
Facebook: [@OPMglobal](https://www.facebook.com/OPMglobal)
YouTube: [@OPMglobal](https://www.youtube.com/OPMglobal)
LinkedIn: [@OPMglobal](https://www.linkedin.com/company/OPMglobal)

Executive summary

Introduction

The Girls' Education Challenge (GEC) was launched by the UK's Department for International Development in 2012 as a 12 year commitment to reach the most marginalised girls in the world and is the largest global fund dedicated to girls' education. The UK is committed to ensuring millions of girls in some of the poorest countries, including girls who have disabilities or are at risk of being left behind, receive a quality education. Through the GEC, we aim to transform the lives of over one million of the world's most marginalised girls through quality education and learning. Access to a good quality education and learning opportunities will empower these girls to secure a better future for themselves, their families and their communities. The first phase of the GEC (2012 - 2017) directly provided quality education for over a million marginalised girls. The GEC is now in its second phase: working through 41 projects in 17 countries. The GEC-Transition Phase II projects are supporting GEC beneficiaries from Phase I to complete primary school, transition to secondary education, and progress on to technical vocational training or employment. Within the second phase, a second cohort of girls are also being supported through the Leave No Girl Behind funding window, which consists of interventions for highly marginalised, adolescent girls who are out of school - either because they have never attended school or have dropped out without gaining a basic education¹.

This paper is based on the combined analysis of baseline data from the 27 GEC Transition (GEC-T) projects conducted by Oxford Policy Management (OPM) and Oxford MeasurEd. This paper explores the learning levels of girls supported by the GEC-T projects (referred to within the report as GEC-T girls). The main contribution of this paper is the identification of the predictors of learning for marginalised girls (using a GEC-T dataset). The paper considers the implications for DFID and GEC. The analysis is research-focused and is not an evaluation of the GEC-T programme. Data from the 15 GEC Leave No Girl Behind (LNGB) projects, which focus on supporting the most marginalised girls, are not included within this analysis.

Methodology

The approach used for this analysis acknowledges the overarching set of social and economic factors that influence schools, teachers, and individuals.

To explore the learning levels and predictors of learning both globally and for marginalised girls, data on global learning levels were downloaded from the United Nations Educational and Scientific and Cultural Organization (UNESCO) Institute of Statistics (UIS) database on the percentage of learners achieving minimum proficiency levels across countries, and Programme for International Student Assessment (PISA)-D data were used to supplement gaps where possible; measures of literacy and numeracy were developed for each GEC-T project within the master dataset; benchmarks based on

¹ DFID's Girls' Education Challenge Fund, (DFID, 2020): www.girlseducationchallenge.org

the global SDG4 indicators '*percent of students achieving minimum level of proficiency in mathematics*' and '*percent of students achieving minimum level of proficiency in reading*' were developed; a literature review of the global literature and evidence base on the predictors of learning across contexts was undertaken; linear regressions were run to identify predictors of learning; and the implications of these findings for DFID and the GEC programme were drawn out from the analysis.

Limitations

Limitations to this analysis centre around the comparability of reading across contexts and the comparability of reading and mathematics benchmarks. Given these difficulties, the major contribution of this paper is not in the presentation of the global learning levels of marginalised girls. Rather, the paper contributes evidence on the predictive effects of selected variables on learning outcomes for marginalised girls and the diversity of these predictive effects across diverse contexts.

Findings

The low levels of learning proficiency in Africa and Asia reflected in the global data are also reflected in the GEC-T data. A large minority of the GEC-T girls did not meet either the reading or mathematics benchmark at baseline. About a quarter met the mathematics benchmark but not the reading benchmark, and approximately 9% met the reading benchmark but not the mathematics benchmark.

The analysis of predictors of learning (using the developed benchmarks across the whole GEC-T portfolio) finds that not speaking the language of instruction at home, the teacher using shouting as a form of discipline, and the primary caregiver having no education negatively predicts meeting a minimum mathematics proficiency level.

What is a predictor of learning?

In statistics, a predictor is a variable that can be used to predict the value of another variable. In this case, that other value is learning.

The predictor could be a feature of the individual learner, of what happens in their school, or of characteristics of their home.

Predictors can be negative or positive. Negative predictors are characteristics or behaviours that are associated with lower learning levels. Positive predictors are characteristics or behaviours associated with higher learning levels.

Age and being overage for grade positively predict meeting the minimum mathematics proficiency benchmark for marginalised girls. This is likely to be explained by the cognitive development of older girls and the low benchmark used for the learning variable. **Across the global GEC-T portfolio, girls reporting that chores interrupt their schooling strongly negatively predicts learning for reading.** As with mathematics, age and grade status positively predict meeting minimum reading proficiency levels and this is likely due to the cognitive development of older girls and the low benchmark used in the regressions.

The teacher using detention and physical punishment is negatively associated with learning.

At the country level, physical, social, economic, linguistic, and teacher behavioural factors emerge as predictors of learning. While the country-level analysis shows a huge amount of variance in the extent to which some characteristics and behaviours predict proficiency across contexts,² some constraints remain. Functional difficulty negatively predicts learning in reading and/or mathematics across several country contexts. Facing multiple economic barriers negatively predicts learning in Tanzania and Afghanistan. Speaking a language other than the language of instruction negatively predicts reading in Afghanistan.

The teacher using physical punishment as a form of discipline only negatively predicts learning, while other teacher behaviours provide mixed results.

At the project level, **an important relationship emerges between learning and poverty** (the role of the primary caregiver and economic barriers), **gendered social norms** (the domestic burden on girls), **disability, and the use of physical punishment in schools.** These relationships remain important across the projects.³

The role of the primary caregiver emerged as significant, as the primary caregiver having no education negatively predicts reading across more projects than the household head having no education. Girls reporting that chores affect school negatively predicts learning.⁴ Reporting at least one functional difficulty⁵ negatively predicts reading and mathematics in both projects with the highest and lowest proportions of girls with disabilities. Reporting an increased number of economic barriers at the household level negatively predicts both mathematics and reading in several projects.

For reading, physical punishment negatively predicts learning. For mathematics, shouting, physical punishment, and detention all negatively predict learning.

Implications

This report emphasises the complexities in the relationships between predictors of learning and learning outcomes. These are relationships of association and not cause. Moreover, not all predictors of learning are amenable to education policy or programme intervention. Therefore, not all negative predictors of learning can be addressed

² Regressions by country reveal mixed results regarding the extent to which the number of reasons travel to school is unsafe and age predict reading outcomes. In four countries, being overage for grade predicted higher reading proficiency, but in two of those countries, age also predicted higher reading outcomes. It is very difficult to disentangle these factors and it is likely that, where both age and overage status predict learning, age is mainly what is being picked up in the prediction. The result of the number of reasons travel to school is unsafe both positively and negatively predicting learning is unexpected. A potential explanation could be the higher baseline learning levels of some projects in FCAS settings. For mathematics, the relationship between age and being overage for grade is again reflected in the results. In Somalia, Sierra Leone, Mozambique, and Kenya, age and/or overage status negatively predicts learning.

³ Despite a huge amount of variance in the extent to which other characteristics and behaviours predict reading and mathematics proficiency across projects is observable. Again age and overage status predict learning in both positive and negative directions in different projects for reading and mathematics.

⁴ In reading.

⁵ Functional difficulty is a term used within the Washington Group Questions: (Washington Group on Disability Statistics, 2020) www.washingtongroup-disability.com/methodology-and-research/conceptual-framework/

by an educational intervention. However, schools are useful sites for intervention and can play an equalising role (Rolleston, 2014).

Based on the analysis of GEC-T baseline learning data, implications for GEC-T emerge as follows.

- **Projects should sensitise communities on gendered social norms and education, but should also provide targeted support to girls with high domestic burdens.** Sensitising communities on education and gendered social norms has been a focus of many education programmes globally, and GEC-T projects should continue work in this area. However, seeing change in gendered norms as a result of intervention is a long and slow process; the provision of support to girls who manage the competing demands of work and school should therefore be provided alongside social norm programmes.
- **Projects should not only focus on eliminating physical punishment in schools, but should also focus on building the capacity of teachers to develop alternative strategies.**
- **The other paper in this published series reported a low prevalence of girls with disabilities within the GEC-T portfolio.** This was hypothesised to have happened due to two principal barriers to projects being inclusive for girls with disabilities. The first was the inability of the government schooling system to meet the needs of girls with disabilities, as most projects reach girls through schools; and the second was the unseen nature of disability and difficulties in finding girls with disabilities due to social stigma and social norms. This paper highlights that, despite the difficulties in reaching girls with disabilities in the GEC-T portfolio, across a number of projects with both low and high percentages of girls with disabilities, disability negatively predicts learning. This is an opportunity for DFID and GEC-T projects to ensure a clear strategy and focus towards further reaching girls with disabilities for support in learning. The success of this strategy could be analysed by conducting a similar analysis of the endline GEC-T data.

Further research

The findings presented in this paper suggest avenues for further research, either using the master dataset developed as part of this project, using the wider evidence base, or by collecting new qualitative and quantitative data, as follows:

1. A full benchmarking activity, qualitatively aligning the difficulty of all subtasks across projects and establishing robust comparisons across the GEC-T baseline assessments, could provide a stronger comparison on the baseline learning levels of GEC-T girls than what has been possible in this project. However, the purpose of such an exercise would need to be clear, as simply comparing learning levels across projects that are not representative of a broader population group is not useful in and of itself.

2. A review of the evidence regarding how initiatives have provided support to girls to manage the competing demands of work and school and 'what works' could provide useful information for GEC-T projects moving forward.
3. Thematic, site-specific qualitative studies on identifying the root causes of physical punishment in schools could provide information to support GEC interventions. This research could include primary qualitative data collection to better understand the conditions under which physical violence in schools takes place and strategies for eliminating violence.

Table of contents

Executive summary	i
List of tables and figures	vii
List of abbreviations	ix
1 Introduction	1
1.1 The GEC-T window.....	1
1.2 Research questions	2
1.3 Theoretical framework and approach	2
1.4 Methodology	4
1.5 Limitations.....	6
2 Starting learning levels and predictors of learning for GEC-T girls.....	9
2.1 Starting learning levels across the GEC-T portfolio	9
2.2 Predictors of learning for marginalised girls across the GEC-T portfolio.....	18
2.3 How do the predictors differ by group and across contexts?	21
3 Summary of the findings and implications for DFID and GEC.....	23
3.1 Summary of findings	23
3.2 Implications for DFID and GEC	23
3.3 Potential further research.....	24
List of references	26
Annex A Learning levels and predictors of learning globally	30
Annex B Multiple linear regression methodology and results	36
Annex C Projects included in proficiency benchmarks	41
Annex D Learning levels by school and individual factors	43

List of tables and figures

Table 1:	Comparing the complexity of texts in an EGRA and a SEGRA reading comprehension task.....	15
Table 2:	First example of a set of EGMA questions and operational requirements...	17
Table 3:	Second example of a set of EGMA questions and operational requirements	17
Table 4:	Individual-, teacher-, and school-level factors to be used in the analysis....	36
Table 5:	Positive and negative relationships between individual and teacher characteristics and meeting the reading benchmark, by country	37
Table 6:	Positive and negative relationships between individual and teacher characteristics and meeting the mathematics benchmark, by country	37
Table 7:	Positive and negative relationships between individual and teacher characteristics and overall achievement in EGRA/SEGRA, by project	38
Table 8:	Positive and negative relationships between individual and teacher characteristics and overall achievement in EGMA/SeGMA, by project	39
Table 9:	Percentage of girls reaching the mathematics benchmark, by project	41
Table 10:	Percentage of girls reaching the reading benchmark, by project	41
Figure 1:	Categories of factors that influence learning outcomes	4
Figure 2:	Percentage of girls who meet the mathematics and reading benchmark	10
Figure 3:	Percentage of girls who meet the reading benchmark, by region	10
Figure 4:	Percentage of girls who meet the mathematics benchmark, by region	11
Figure 5:	Percentage of girls who meet the mathematics benchmark, by country	11
Figure 6:	Percentage of girls who meet the reading benchmark, by country	12
Figure 7:	Percentage of girls who meet the reading benchmark, by FCAS status	13
Figure 8:	Percentage of girls who meet the mathematics benchmark, by FCAS status	13
Figure 9:	Unique contribution of predictors of meeting the mathematics benchmark (standardised beta weight coefficient)	20
Figure 10:	Unique contribution to the prediction of meeting minimum reading benchmarks—GEC-T (standardised beta weight coefficient)	21
Figure 11:	Percentage of children who meet reading benchmarks globally	31
Figure 12:	Percentage of children who meet mathematics benchmarks globally	31
Figure 13:	Average percentage of children achieving minimum proficiency in mathematics, by fragility status	32

Figure 14:	Average percentage of children achieving minimum proficiency in reading, by fragility status	32
Figure 15:	Percentage of girls who meet the mathematics benchmark, by grade	43
Figure 16:	Percentage of girls who meet the reading benchmark, by grade	44
Figure 17:	Percentage of girls who meet the reading benchmark, by age	44
Figure 18:	Percentage of girls who meet the mathematics benchmark, by age	45
Figure 19:	Percentage of girls who meet the reading benchmark, by age/grade status	45
Figure 20:	Percentage of girls achieving mathematics benchmark, by age/grade status	46
Figure 21:	Percentage of girls who meet reading and mathematics benchmarks, by whether chores affect their schooling	46
Figure 22:	Percentage of girls who meet reading and mathematics benchmark, by functional difficulty	47
Figure 23:	Percentage of girls who meet reading and mathematics benchmarks, by marital status	47
Figure 24:	Percentage of girls who meet reading and mathematics benchmarks, by number of economic barriers	48

List of abbreviations

DFID	UK Department for International Development
EGMA	Early Grade Mathematics Assessment
EGRA	Early Grade Reading Assessment
FM	Fund Manager
GEC	Girls' Education Challenge
GEC-T	Girls' Transition
LNGB	Leave No Girl Behind
MC	Multi-Country
NGO	Non-Governmental Organisation
OECD	Organization of Economic Cooperation and Development
OPM	Oxford Policy Management
PISA	Programme for International Student Assessment
SACMEQ	Southern and Eastern Africa Consortium for Monitoring Educational Quality
SDG4	Sustainable Development Goal 4
SeGMA	Secondary Grade Mathematics Assessment
SeGRA	Secondary Grade Reading Assessment
SES	Socioeconomic Status
UIS	UNESCO Institute of Statistics
UNESCO	United Nations Educational and Scientific and Cultural Organization

1 Introduction

1.1 The GEC-T window

The UK is committed to ensuring millions of girls in some of the poorest countries, including girls who have disabilities or are at risk of being left behind, receive a quality education. DFID's GEC⁶ supports up to 1.5 million of the world's most marginalised girls to complete a full cycle of either primary or secondary education. Phase I of GEC (2012–2017) was funded with a £355 million investment and targeted 1.4 million marginalised girls through 37 different projects. GEC is now in Phase II (running from 2017 to 2025), during which 41 projects will receive £500 million to support their activities.

GEC's Phase II consists of two funding windows:

1. a GEC-T window to continue funding 27 successful GEC Phase I projects across 15 countries and to ensure that one million marginalised girls transition successfully from primary education into secondary education, further education, vocational education, or training; and
2. an LNGB window to fund 14 targeted 'catch-up' projects for up to 250,000 highly marginalised girls in 10 countries.

GEC projects are designed and delivered by implementing partners, including international NGOs, social enterprises, and private sector organisations, working mostly within government schools. Projects were procured through a challenge fund procurement process. Projects deliver a broad range of interventions, including tailored classroom teaching, teacher development, and school improvement; educational technology and distance learning; community engagement; and financial support to girls, their families, and their schools. All projects have a strong focus on accelerating girls' learning outcomes. GEC-T projects aim to improve girls' literacy and numeracy skills and support them to acquire the relevant knowledge, skills, and attitudes needed for life and work. Projects also aim to reduce school drop-out during adolescence, including tackling harmful social and gender norms, child marriage, early pregnancy, domestic work, or violence.

All GEC-T projects are being evaluated through independent, mixed-methods evaluations. DFID currently holds large baseline datasets and baseline evaluation reports from the 27 independent GEC-T project level evaluations. These contain valuable information on the characteristics of marginalised girls supported by GEC-T projects, their learning levels, and the barriers they face in attending school, performing well, and transitioning into secondary school or vocational/employment pathways. The combined analysis of these datasets has the potential to contribute significantly to the evidence base on girls' marginalisation.

This paper is based on the combined analysis of the 27 GEC-T baseline datasets conducted by OPM and Oxford MeasurEd. The analysis of this data helps to understand who the GEC-T girls are and to describe their levels of marginalisation; provides a deeper

⁶ <https://www.gov.uk/guidance/girls-education-challenge>

understanding of the safety and security challenges in GEC-T contexts; addresses the dearth of evidence on the learning levels of marginalised girls in diverse contexts; provides evidence to support DFID and the GEC Fund Manager (FM) to make informed decisions about how to provide appropriate technical advice and support across the portfolio of GEC programmes; provides evidence to inform DFID's global influencing and communication strategies; and supports the accountability of public funds.

This paper explores the learning levels of GEC-T girls. The paper reviews the literature and evidence on the learning levels and predictors of learning globally and explores contextual and regional differences. The paper presents findings on the baseline learning levels of GEC-T girls and the predictors of learning, exploring contextual and regional differences. The paper explores the extent to which the predictors of learning are similar or different for GEC-T girls in comparison with other populations and considers the implications for DFID and GEC. It is outside of the scope of this paper to assess the GEC-T programmes or programme Theories of Change.

1.2 Research questions

As agreed at inception and documented in the Scoping Paper for this research, this paper provides answers to a set of questions using a range of data sources, including the GEC-T master dataset as a sample of data available on marginalised girls. These questions include the following.

1. What are learning levels across low- and middle-income countries?
2. How do these differ by region?
3. What are the predictors of learning levels across contexts and social groups?
4. What are the starting learning levels (literacy and numeracy) of the GEC-T girls?
5. How do the learning levels differ by individual and contextual characteristics?
6. What are the key predictors of learning outcomes for marginalised girls?
7. What are the implications for DFID and the GEC programme?

At the request of DFID, responses to research questions 4, 6, and 7 are presented in the body of this report, and responses to questions 1, 2, 3, and 5 are included as Annexes A and D.

It is important to note that the purpose of this analysis is to use the GEC-T baseline datasets to better understand the predictors of learning for marginalised girls. As the analysis uses baseline data, the analysis is not focused on comparing learning outcomes across GEC-T projects or attribute learning levels to GEC-T. The analysis focuses on researching the learning predictors for marginalised girls, not on evaluating the GEC programme.

1.3 Theoretical framework and approach

Learning outcomes are the result of a complex set of factors and influences both observable and unobservable, including (and not limited to) wider societal and national factors (Branigan, McCallum, and Freese 2013), the schooling system (OECD, 2003), the

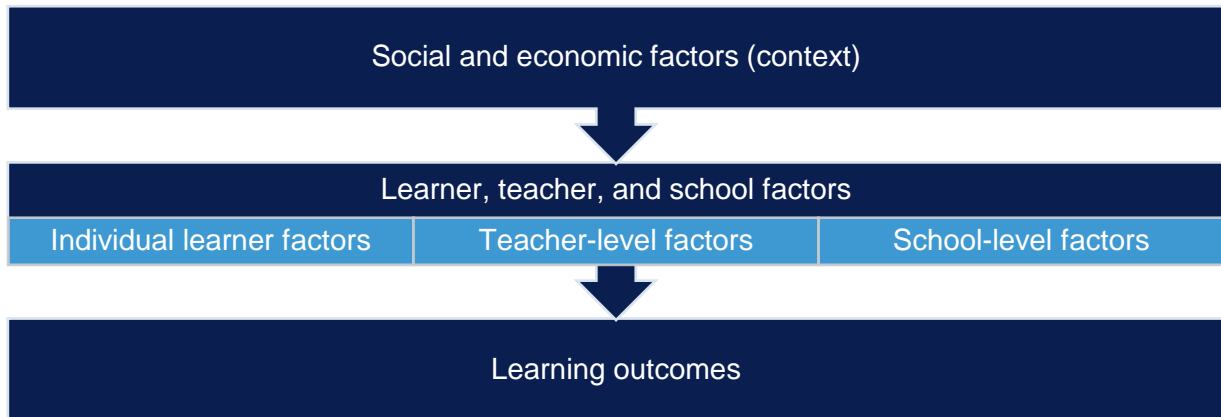
characteristics of schools and teachers (both inputs and practices), and the characteristics, attitudes, behaviours, and practices of learners (Hattie, 2009). The ways in which these factors operate and interact with each other are complex and can never be fully mapped or accounted for in an analytical set-up. In addition, some of these factors are more amenable to measurement (and intervention) than others.

Societal factors, national planning, and programme implementation can play an important role in improving learning outcomes (Bermeo, 2014). The median level of achievement in many low- and middle-income countries is approximately the fifth percentile of the distribution in Organization of Economic Cooperation and Development (OECD) countries (Couch and Rolleston, 2017). In addition, social and sociolinguistic factors are especially important for language and literacy outcomes.

At the school level, factors that affect learning outcomes and learning progress include the infrastructure within the school, teaching and learning resources, school leadership, and the learning culture within the school (Hattie, 2009). The quality of teaching, collaboration, and professional support between teachers; teacher classroom attendance; and teacher feedback are examples of teacher-level factors that impact on learning outcomes (Hattie, 2009). The effects of these factors are difficult to identify and quantify precisely in empirical analysis, not least because they vary widely across education systems and contexts.

In the global literature, within countries, the most powerful predictors of learning outcomes are what learners themselves bring with them into the classroom, across geographical and schooling contexts. Intellectual, social, and cultural factors, along with household-specific factors, significantly predict pupil performance in most standardised learning assessments (Hattie, 2009; Outhred and Beavis, 2012; Outhred and Beavis, 2013; Dahl and Lochner, 2005). The international literature is broadly in agreement that family background factors powerfully influence student learning outcomes. Individual factors include household wealth and parents' education, gender, language background, and age.

In line with the evidence base, our analytical approach acknowledges the overarching set of social and economic factors that influence schools, teachers, and individuals. Our analytical approach categorises individual-, teacher-, and school-level factors as variables that influence learning outcomes, as described in Figure 1. However, very few school-level factors were available in the data, and this is a limitation of the analysis.

Figure 1: Categories of factors that influence learning outcomes

1.4 Methodology

To answer the research questions above, the following methodology was employed.

1. Data on global learning levels were downloaded from two principal sources.

- a. The UIS has been given the mandate to ‘work with partners to develop new indicators, statistical approaches and monitoring tools to better assess progress across the targets related to UNESCO’s mandate, working in coordination with the Education 2030 Steering Committee.’ This has included establishing the Technical Cooperation Group to build consensus on the Sustainable Development Goal 4 (SDG4) measurement agenda, the Global Alliance to Monitor Learning to provide concrete solutions to develop new indicators and set standards for good practices in learning assessment, and the Inter-Agency Group on Education Inequality Indicators to set the standards across the United Nations system to report and interpret household survey data, especially in relation to measuring indicators on equity.

Based on this mandate, UIS has developed a global database on the percentage of learners achieving minimum proficiency levels across countries. These data were downloaded from the UIS data centre to be used to assess global learning levels.

- b. For countries that participate in PISA-D, data was added using the lowest PISA-D proficiency benchmark as the definition of ‘a minimum level of proficiency’.

These global learning levels are presented in Annex A.1.

1. A literature review of the global literature and evidence base on the predictors of learning was undertaken.

Data from international and national surveys, grey literature, and academic publications were reviewed. The predictors of learning were identified and described by context.

The global literature review on the predictors of learning is presented in Annex A.2.

2. Measures of literacy and numeracy were developed for each GEC-T project within the master dataset.

Data collected by project evaluators as part of the baseline evaluations across the 27 GEC-T projects were merged into one master dataset. The dataset includes variables that report on individual-, household-, and community-level factors, teacher-level factors, and school-level factors, where these are comparable across projects.⁷

To assess literacy levels, all project evaluators were guided to follow a similar structure to an Early Grade Reading Assessment (EGRA) and to adopt a new testing framework, referred to as the GEC Secondary Grade Reading Assessment (GEC SeGRA). Girls were assessed on a number of tasks including letter knowledge, phonemic awareness, phonics, decoding, vocabulary, reading fluency, and comprehension. Reading fluency was assessed through a timed fluency sub-task, with early grades using less complicated passages and higher grades using more complicated passages with more challenging comprehension questions. There are limitations to this approach to assessing literacy levels; these have been outlined in the Scoping Paper for this paper series and they are summarised in the limitations section of this report.

To assess numeracy levels, all project evaluators were guided to follow a similar structure to an Early Grade Mathematics Assessment (EGMA) and to adopt a new framework referred to as the GEC Secondary Grade Mathematics Assessment (GEC SeGMA). Girls were assessed on a number of tasks the identification and manipulation of numbers, including addition, subtraction, multiplication, and division. The assessment finishes with a number of word problems. In higher grades, space data, algebra, and more complicated word problems were included.

An overall aggregate learning measure for reading and mathematics was created by averaging the scores from each task in the assessment. All tasks were weighted equally.⁸

2. Learning benchmarks for minimum proficiency levels in reading and mathematics were developed and applied to the literacy and numeracy measures of the GEC-T projects. These benchmarks were based on the global SDG4 indicators ‘percentage of students achieving a minimum level of proficiency in mathematics’ and ‘percentage of students achieving a minimum level of proficiency in reading’

A minimum level of proficiency in mathematics was defined as being able to respond correctly to at least half of the EGMA word problems. This benchmark was chosen as

⁷ Each project administered questionnaires with a combination of items required by the FM and items specific to the individual project evaluation. Only items common across all projects were included as predictors. While some items on background characteristics are common across projects, they may not always be entirely comparable as they are open to interpretation or contextual differences. For example, any measure of financial constraints is informed by norm-referenced ideas of what ‘financial difficulty’ means in each context, and may not strictly be comparable between countries.

⁸ The exact questions asked and the level of difficulty of the questions differs across projects. It is therefore not possible to use the aggregate learning measure to directly compare performance of girls across projects.

solving a mathematical problem represents the domain of mathematics better than other EGMA subtasks.

A minimum level of proficiency in reading was defined as being able to respond correctly to at least half of the EGRA (first option) or SEGRA (where EGRA was not available) reading comprehension questions. The reading benchmark was chosen as it best represents reading for meaning. Understanding the meaning behind texts is representative of the domain of reading to a greater extent than comparing the number of words a child can read, particularly because learning assessments in GEC are carried out in a variety of languages and reading speeds differ by language. This makes it more problematic to choose reading fluency / speed as a reading benchmark compared to reading comprehension.

3. Relationships between reaching reading and mathematics benchmarks and individual, school, and country characteristics were analysed across the GEC-T portfolio.

Relationships between reaching the reading and mathematics benchmarks across the GEC-T portfolio and the background characteristics of the GEC-T girls were explored through descriptive statistics (presented in Annex D) and through the use of multivariate linear regression models.

4. Relationships between learning levels and individual and school characteristics were analysed within individual GEC-T projects.

Within each individual project dataset, linear regression models were run using mathematics and reading total scores as the dependent variable to identify predictors of learning within each project. Predictors of learning were compared across projects and patterns were identified.

Similarities and differences in the patterns predictors of learning were identified. The implications of these findings for DFID and the GEC programme were drawn out from the analysis.

What is a predictor of learning?

In statistics, a predictor is a variable that can be used to predict the value of another variable. In this case, that other value is learning.

The predictor could be a feature of the individual learner, of what happens in their school, or of characteristics of their home.

Predictors can be negative or positive. Negative predictors are characteristics or behaviours that are associated with lower learning levels. Positive predictors are characteristics or behaviours associated with higher learning levels.

1.5 Limitations

Limitations to this analysis centre around the comparability of measures of reading across contexts and the comparability of reading and mathematics benchmarks.

1. Global indicators are open to interpretation, as are most education benchmarks that are not derived from specific high-quality assessment programmes (PISA, the Southern and Eastern Africa Consortium for Monitoring Educational Quality

(SACMEQ), Trends in International Mathematics and Science Study, and Progress in International Reading Literacy Study). It is also important to note that the purpose of GEC-T assessments is to track progress in learning outcomes over time for each individual project. Project evaluators were asked to use texts from an appropriate grade that they were testing. The assessments were not designed to be comparable across the portfolio.

As a result, the difficulty of the problems and questions put to learners across the portfolio will differ. The benchmark is designed for learners who have completed two or three years of schooling, while many GEC-T girls have completed more than three years. However, the ability to solve mathematical problems and understand the meaning of text are important indicators of proficiency. While there are clear limitations in comparing learning levels using data not designed for comparison, these are the same challenges faced by the international community in monitoring progress against SDG4. Therefore, while this analysis suffers from the same comparability challenges experienced globally, adhering to the same minimum proficiency benchmarks as the global community is a useful basic comparer across the GEC-T portfolio of projects.

The global benchmark of '*minimum proficiency*' is a pragmatic solution to the need to monitor learning outcomes globally, in the absence of globally comparable learning measures. The agreed global indicator for reading, for example, refers to 'a grade 2 or 3 level of text'. While this seems a fairly straightforward description of the benchmark, there are insurmountable technical difficulties with global comparisons based on these simplistic benchmarks.

These difficulties apply to the GEC-T baseline data no more or less than any other dataset using scores from different assessments. Given the intractable technical issues regarding comparability and global learning benchmarks, this paper takes a pragmatic approach (as has the UIS, which has the global mandate to monitor global learning).

However, given these difficulties, the major contribution of this paper is not in the presentation of the global learning levels of marginalised girls. Rather, the paper contributes evidence on the predictive effects of selected variables on learning outcomes for marginalised girls and the diversity of these predictive effects across diverse contexts. The major contribution of this paper is the identification of predictors of learning for marginalised girls.

1. Approaches to teaching and phonological properties of a language itself influence scores on reading assessments. This means that comparing data across contexts is often not strictly appropriate, as subtasks such as phonological knowledge may be heavily influenced by pedagogical approaches or the language tested.

For this reason, we use reading comprehension as the benchmark. This is informed by the importance of 'reading for meaning', which is equally important across contexts and languages.

In seven cases, SEGRA (rather than EGRA) reading comprehension results were used for the benchmarking. In theory, SEGRA should be a more difficult assessment than EGRA; however, this paper reports on a comparison of EGRA and SEGRA texts and questions, and it was not clear that SEGRA texts used in this analysis are more difficult.

2. There is a great deal of linguistic diversity in many of the contexts in which GEC-T programmes operate. Learners who are being assessed in a language other than their mother tongue are actually being assessed in part on *language* and not solely on *literacy*. This said, GEC-T assesses girls in the language of instruction and their knowledge of this language is important for future learning to take place. Therefore, knowledge of the language of instruction is an important measure of learning.
3. Associations between reading benchmarks and individual, school, and country contexts are purely descriptive. These relationships are not causal; however, they provide useful information on the learning and individual profiles of GEC-T girls.
4. This report analyses the relationships between sets of numbers (background characteristics and learning levels) and does not provide extensive explanations for the findings. Contextual and qualitative analysis could provide further explanation regarding why the relationships between these numbers exist.

2 Starting learning levels and predictors of learning for GEC-T girls

There is a dearth of high-quality data available on the learning outcomes of highly marginalised young people. Even less data are available in marginalised settings on the individual- and community-, teacher-, and school-level characteristics associated with attendance, transition, and learning that can be acted upon by policymakers or through programme intervention.

In this section of the report, we focus on the learning levels of girls within the GEC-T portfolio. The starting learning levels of the girls are presented, using the percentage of girls reaching minimum learning levels. In Section 2.2, the results of a series of linear regressions are presented. The predictors of learning are presented based on the achievement of minimum reading and mathematics proficiency benchmarks across the whole GEC-T portfolio, the achievement of minimum reading and mathematics proficiency benchmarks for specific countries (where available), and GEC-T girls scores on specific assessments by project. The similarities and differences in the predictors of learning across countries and projects are described.

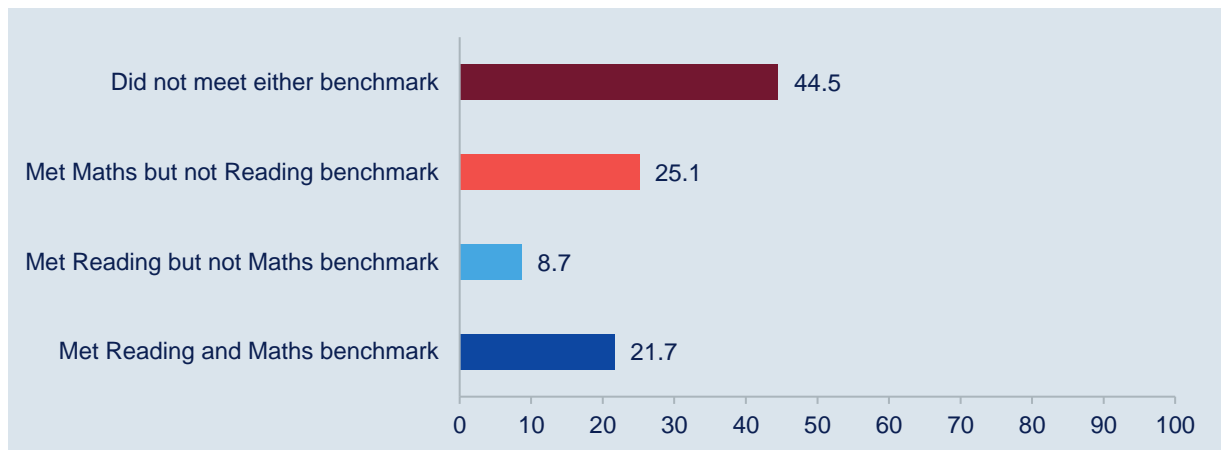
In interpreting the results, it is important to emphasise that these relationships are not causal, but rather provide a picture of the connections between certain factors and the likelihood of GEC-T girls meeting minimum benchmarks at baseline (at the portfolio and country levels) or scoring better on reading and mathematics assessments (projects).

2.1 Starting learning levels across the GEC-T portfolio

Overall, 48.1% of GEC-T girls were meeting the minimum mathematics benchmark and 38.6% were meeting the minimum reading benchmark when they first engaged with a GEC-T project.⁹ Approximately 22% of girls met both the mathematics and the reading minimum benchmark.

A large minority of the GEC-T girls did not meet either the reading or mathematics benchmark at baseline. About a quarter met the mathematics benchmark, but not the reading benchmark, and approximately 9% met the reading benchmark but not the mathematics benchmark.

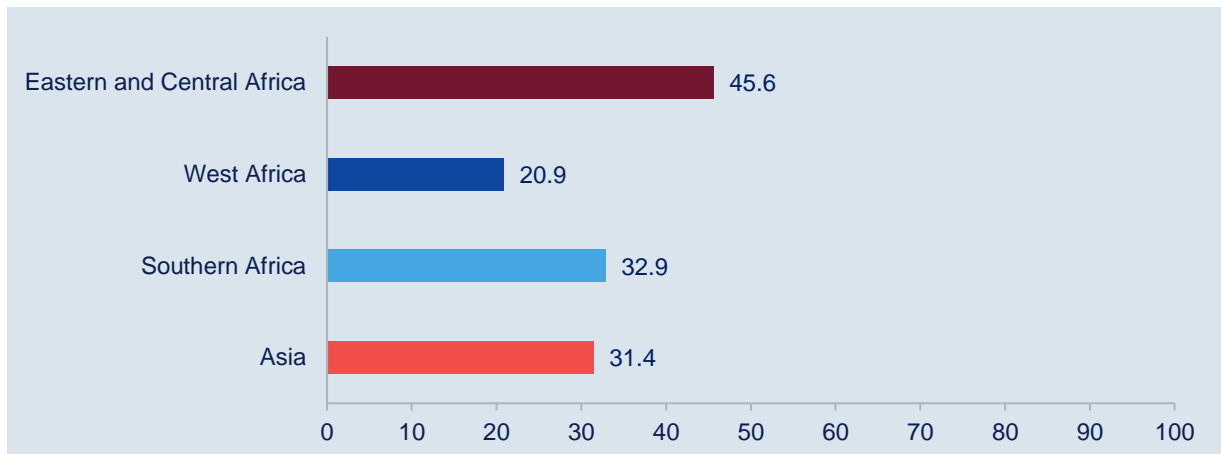
⁹ Reading and mathematics minimum proficiency benchmarks are not comparable. It is not possible to conclude whether GEC-T girls are better at mathematics than at reading, for example, as it is not known if the minimum proficiency range for mathematics is more or less difficult than for it is for reading.

Figure 2: Percentage of girls who meet the mathematics and reading benchmark

Source: GEC-T master dataset

2.1.1 Country context for GEC-T girls

The percentage of GEC-T girls who meet the reading and mathematics benchmarks at baseline differs by region¹⁰. The percentage of girls who meet the reading benchmark is higher in Eastern and Central Africa compared to the other regions. About the same percentage of girls meet the reading benchmark in Southern Africa and Asia, and the lowest percentage of girls achieve the reading benchmark in West Africa.

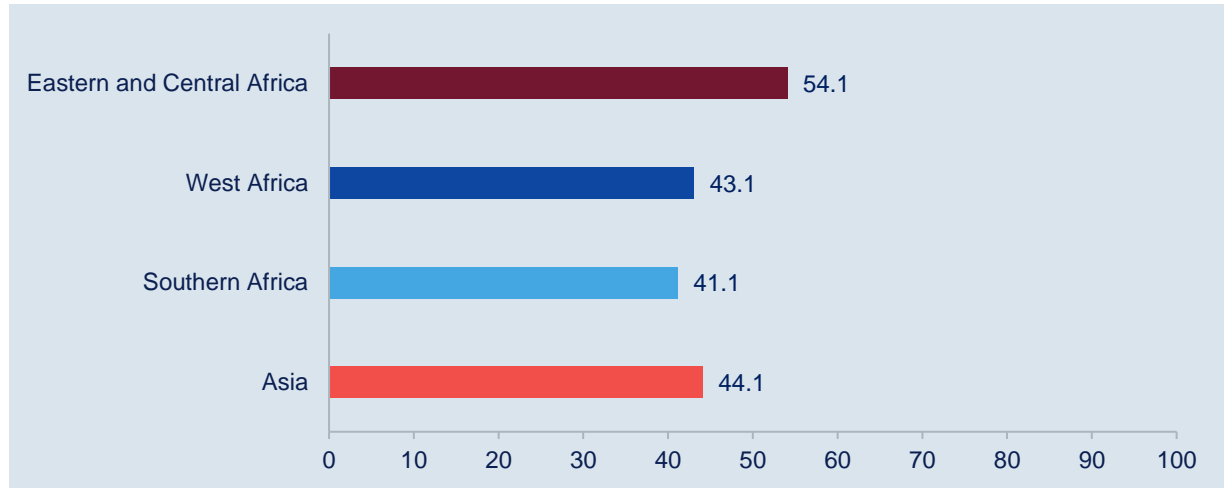
Figure 3: Percentage of girls who meet the reading benchmark, by region

Source: GEC-T master dataset

¹⁰ Eastern and Central Africa includes the DRC, Ethiopia, Kenya, Rwanda, Somalia, Tanzania, and Uganda. There are 18 projects operating in this region. West Africa includes Ghana, Sierra Leone, and Nigeria. There are five projects operating in this region. Southern Africa includes Mozambique, Zambia, and Zimbabwe. There are four projects operating in this region. Asia includes Afghanistan and Nepal. There are four projects operating in this region.

As with the reading benchmark, the percentage of girls who meet the mathematics benchmark is highest in Eastern and Central Africa. Approximately the same percentage of girls meet the mathematics benchmark in West Africa, Southern Africa and Asia.

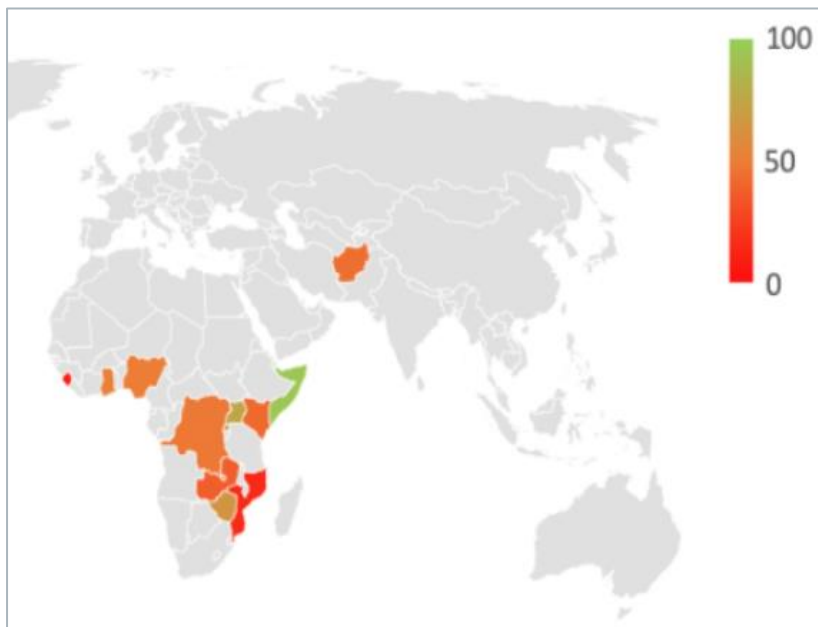
Figure 4: Percentage of girls who meet the mathematics benchmark, by region



Source: GEC-T master dataset

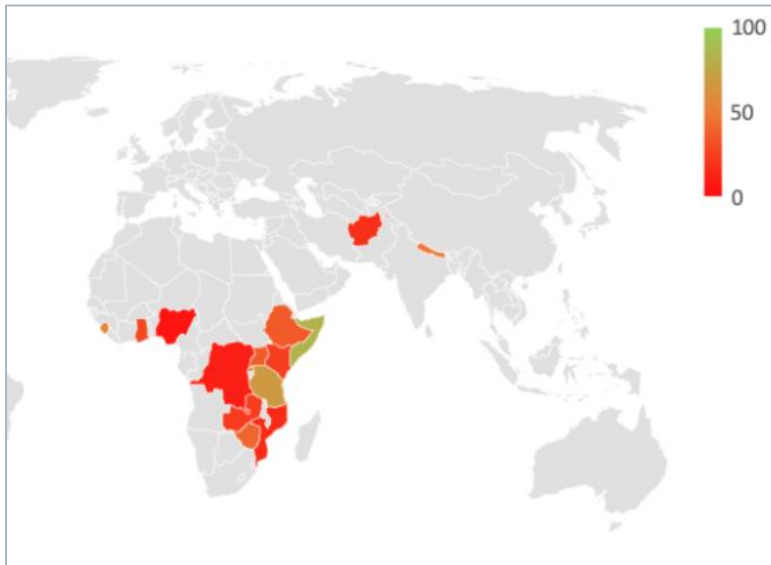
At the country level, the only country where a large majority of girls meet the mathematics benchmark is Somalia. The percentage of girls who meet the mathematics benchmark at baseline is particularly low across all other countries, with the exception of Kenya.

Figure 5: Percentage of girls who meet the mathematics benchmark, by country



Source: GEC-T master dataset

The percentage of girls who meet the reading benchmark is also high in Somalia. The percentage of girls who meet the reading benchmark is particularly low in Mozambique. Additional primary research would be necessary to explore why this is the case and understand implications of these findings.

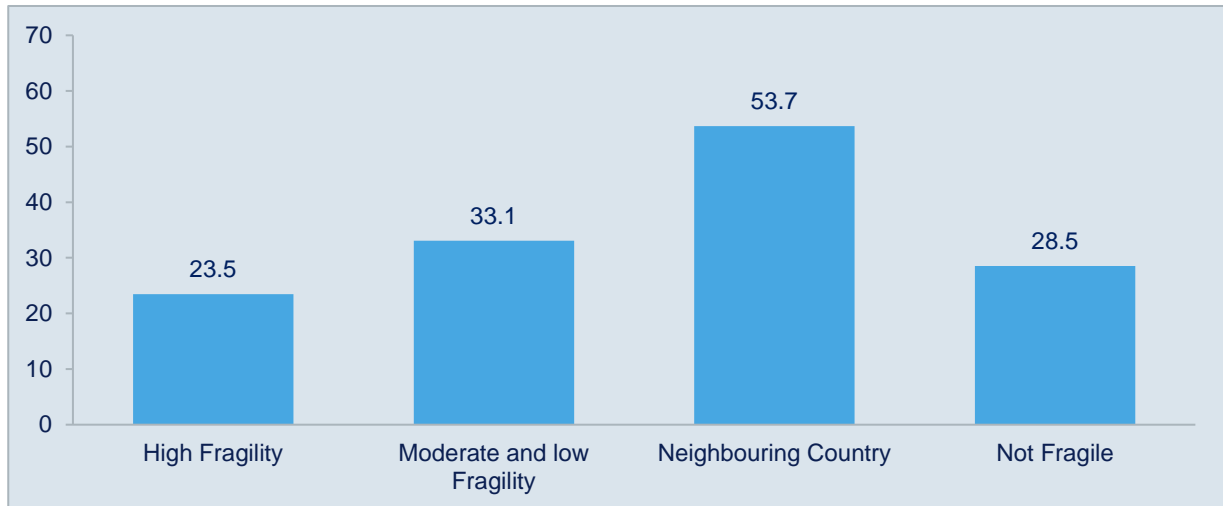
Figure 6: Percentage of girls who meet the reading benchmark, by country

Source: GEC-T master dataset

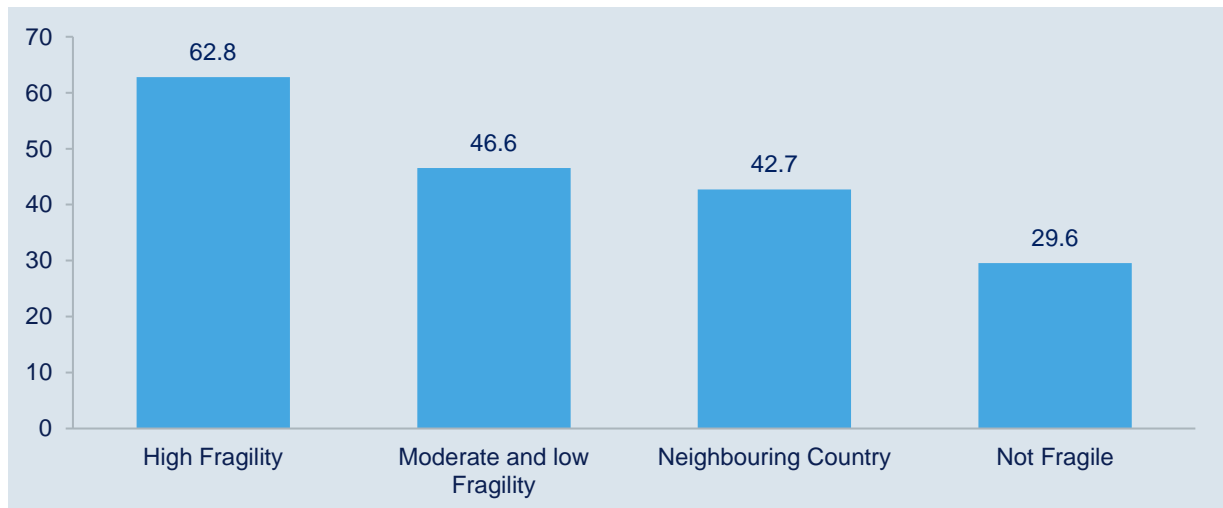
GEC-T projects operate across different contexts of fragility and conflict. DFID's FCAS ranking classifies countries based on their level of instability, insecurity, and conflict. Countries are classified as high fragility, moderate fragility, low fragility, and neighbouring country (and non-fragile countries that are not included on the list). Neighbouring countries border at least one country classified as 'high fragility'.

GEC-T girls in high fragility contexts are less likely than those in other settings to meet the reading benchmark. However, they are more likely to meet the mathematics benchmark than those in other settings.

The pattern of the percentage of GEC-T girls meeting minimum benchmarks for reading and mathematics and fragility status are almost opposite. For reading, as fragility decreases, the percentage of girls meeting the benchmark is higher, with the exception of non-fragile countries. For mathematics, as fragility decreases, the percentage of girls meeting the benchmark is lower. Nothing in the data or project evaluation reports refers to or explains this observation.

Figure 7: Percentage of girls who meet the reading benchmark, by FCAS status

Source: GEC-T master dataset

Figure 8: Percentage of girls who meet the mathematics benchmark, by FCAS status

Source: GEC-T master dataset

2.1.2 What GEC-T girls know and can do at baseline

As has been noted earlier in this report, there are large complexities in comparing scores across different tests. However, when reviewing what GEC-T girls know and can do at baseline, the application of scoring at least 50% on reading comprehension and at least 50% on mathematics word problems reasonably represents ‘a minimum level of proficiency’.

For reading comprehension, the linguistic complexity of the text, the complexity of the language of administration, and the demands of the cognitive processes required to respond to different comprehension questions collectively determine the difficulty of the task.

The question of linguistic complexity is not easy and linguists have developed various frameworks to define it. The mode (formal versus informal; planned or

spontaneous), tenor (tone based on the relationship with the audience), and field (specialised versus common sense/lay) of a piece of text determine the 'complexity' of a text (Leung, 2016). Quantitative frameworks to measure the complexity of text include counting the features of the textual elements of words, syntax, or discourse—for example, word frequency, sentence length, number of pronouns, cohesive devices, and genre markers (Cunningham & Mesmer, 2014).

To add to the complexity of the task, languages themselves differ in complexity.

Efforts to measure the difficulty of languages have included counting the number of points in 'a typical sentence' that are capable of receiving inflection (Nichols, 1986) and measuring the length of the grammar that a language requires (McWhorter, 2001).

Finally, not only does the complexity of text determine the difficulty of reading comprehension items, but the underlying cognitive machinery processes required to respond to different questions also determines the difficulty of reading comprehension items (Juola).

Table 1 compares the linguistic complexity and the requirements of the comprehension questions for an EGRA and SEGRA set of reading comprehension tasks within the GEC-T portfolio.

The EGRA text is more complex in terms of tone, sentence length, and the use of abstract nouns. The SEGRA text is more complex in terms of length. Both texts are basic in terms of their content. They present information or convey a simple story and are predominantly descriptive. There is no play on words or demand for the reader to read for inferred meaning, appreciative reactions, or critical evaluation. There is no demand for the reader to produce fresh or original ideas not explicitly stated in the reading.

The set of questions developed to assess comprehension in both EGRA and SeGMA predominantly rely on recalling read words. However, SEGRA does include one question that could be interpreted as a task to infer meaning and one question that requires identification of a theme or narrative in the text. However, it is not clear if the marking matrix required the learner to correctly identify a theme or narrative in the story to gain a 'correct' score on this item.

Table 1: Comparing the complexity of texts in an EGRA and a SEGRA reading comprehension task

Text examples from two different projects	Text complexity description	Questions	Requirements of the comprehension questions
<p>EGRA example¹¹</p> <p>Our beloved homeland is a mountainous country. Most of the mountains are located in the central, east and other parts which has a good effect on the climate of our country. Wide ranges of icebergs are made due to snow fall in these mountains; that is why Large Pamir of Badakhshan is considered a Natural Park. The ranges of Hindustan, Solomon and Turkistan mountains are also worth mentioning for having massive and precious mines. Additionally, existing Lakes here also beautify the nature of Afghanistan, like Ghazni Lake and Amir Dam in Bamyam which have grabbed foreigners' attention. Forests are also another part of natural resources.</p>	<p>This reading passage is shorter than the SEGRA reading passage. It is 105 words long. It includes fewer sentences (six), but the average sentence length is longer (17.5 words).</p> <p>The mode of the text used in the assessment includes the presentation of the text, its representations, and any technical or contextual knowledge.</p> <p>The text is presented to the student using large (16-point) Calibri font. There are no representations of speech present in the text. The text is descriptive of a region in Afghanistan and presents geographical and geosocial information.</p> <p>The lexis is formal and includes subject-specific lexical sets in the form of geographical knowledge.</p> <p>The text predominantly uses concrete nouns, but many abstract nouns are also present.</p> <p>Likewise, the text predominantly uses stative, rather than active, verbs.</p> <p>There is no use of euphemisms, ludic language, or puns in the text.</p>	<p>Where are most of the mountains located? (Answer: central and east regions)</p> <p>What part of Badkahshan is famous for its icebergs? (Answer: Large Pamir)</p> <p>What can be found in the Hindustan mountains? (Answer: massive and precious mines)</p> <p>Can you name a lake that has attracted attention from foreigners? (Answer: Ghazni Lake, or Amir Dam in Bamyam)</p> <p>What other type of natural resource is mentioned in the text? (Answer: Forests)</p>	<p>Recall words read in the text.</p> <p>Recall words read in the text.</p> <p>Comprehend and interpret text using different words.</p> <p>Comprehend and interpret text using different words.</p> <p>Recall words read in the text.</p>
<p>SEGRA example</p> <p>Mr. and Mrs. Tumaini lived in a village called Mwaze. They lived happily with their two children Imani and Lulu. Imani was their elder daughter while Lulu was</p>	<p>This reading passage is longer than the EGRA example. It is made up on 11 sentences, but the average sentence length is smaller (15.75).</p>	<p>Where did Mr and Mrs. Tumaini live?</p>	<p>Recall words read.</p> <p>Recall words read in the text.</p>

¹¹ These examples were not necessarily administered in English, but have been translated to English and used as an example of text content for an English readership. The difficulties with comparing texts across languages are well-documented in this paper and in the Scoping Paper for this assignment.

Text examples from two different projects	Text complexity description	Questions	Requirements of the comprehension questions
<p>their youngest daughter. Imani was very intelligent and managed to pass her Standard Seven examinations and went to secondary school.</p> <p>After completing her studies, Imani went to town and got a good job. She was paid well and so she managed to build a good house for her parents and younger sister. After she had worked for a few years, she called her father and told him that she would send them money which they should use to buy chickens (layers). When Imani's father told his wife, she was very happy and she immediately started preparing a place to build the poultry house.</p> <p>When Imani sent the money, they bought the chicks and were able to keep them well in a nice poultry house. The chickens lay eggs and they sell them at the market. Mr and Mrs. Tumaini have a large poultry house. They have built two other houses for rent by using the money they get from the eggs.</p>	<p>The mode of the text used in the assessment includes the presentation of the text, its representations, and any technical or contextual knowledge.</p> <p>The text is presented to the student using large (14-point) Times New Roman font. There are no representations of speech present in the text. The text is a story about a family and describes a series of events.</p> <p>The lexis is informal and does not rely on subject-specific or technical lexical sets.</p> <p>The text uses concrete nouns and very few abstract nouns.</p> <p>The text uses predominantly stative verbs and few active verbs.</p> <p>There is no use of euphemisms, ludic language, or puns in the text.</p>	<p>How many children did Mr. and Mrs. Tumaini have?</p> <p>Write the names of Mr. and Mrs. Tumaini's children.</p> <p>Which of their children went to secondary school?</p> <p>When did Imani get a job?</p> <p>What two things did Imani do for her parents?</p> <p>How did Mr. and Mrs. Tumaini use the money from selling the eggs?</p> <p>From the passage, how can you tell that Imani's mother was happy?</p> <p>Suggest a suitable heading to the passage.</p>	<p>Writing task (copying) in the reading comprehension sub-task.</p> <p>Recall words read in the text.</p> <p>Recall words read in the text</p> <p>Comprehend and interpret text using different words.</p> <p>Comprehend and interpret text using different words.</p> <p>Infer meaning from text.</p> <p>Identify the theme or narrative of a text.</p>

For mathematics, word problems require learners to undertake operations and more complex word problems require learners to determine the order of operations. The analysis below shows that, while the difficulty of mathematics word problems are different across tests, both examples capture basic mathematical operations and order of operations knowledge.

The 'operations' are addition, subtraction, multiplication, division, and exponentiation. The 'order' of these operations states which operations take precedence before other operations.

Table 2 and Table 3 show that both examples of EGMA questions require learners to undertake basic multiplication, subtraction, and addition operations and understand the order of operations in a least one question. However, the first example includes triple-and double-digit numbers more often, increasing the difficulty of the questions.

Table 2: First example of a set of EGMA questions and operational requirements

First example set of EGMA questions	Description of the mathematics operational requirements
Suppose you went to the market with two AFN 100 notes and five AFN 5 notes. So how much money did you take in total to the market?	Multiplication of triple- and single-digit numbers based on a word problem.
Suppose you save AFN 20 each month. How much will you have saved after 12 months?	Multiplication of two triple-digit numbers based on a word problem.
Suppose you have AFN 240 in total and a chicken costs AFN 80. How many chickens can you buy?	Subtract a double-digit figure from a triple-digit number based on a word problem.
Suppose you take a loan of AFN 100 with an amount of AFN 4 paid monthly for interest. How much money in total will you have to return after one year (with interest and principal amount combined)?	Multiplication of a single-digit and a double-digit number added to a triple-digit number, based on a word problem.
We know that 365 days equal one year. Suppose it rained for 112 days last year. If so, how many days did it not rain last year?	Subtract one triple-digit number from another based on a word problem.

Table 3: Second example of a set of EGMA questions and operational requirements

Second example set of EGMA questions	Description of the mathematics operational requirements
There are 5 children in a house. 3 more children go into the house.	Addition: two single-digit numbers, totalling to a single digit, based on a word problem.
There are 9 children in the house. 3 are boys. The others are girls. How many girls are in the house?	Addition: two single-digit numbers, totalling to a double digit, based on a word problem.
There are 4 children in John's house. There are 7 children in Mary's house. How many more children must go into John's house so that it has the same number of children as Mary's house?	Single-digit number subtracted from a single-digit number, based on a word problem.
There are some children in a house. 3 more children go into the house. Now there are 10 children in the house. How many children were in the house at the beginning?	Single-digit number subtracted from a double-digit number, based on a word problem.

Second example set of EGMA questions	Description of the mathematics operational requirements
There are 12 sweets. 3 children share the sweets equally. How many sweets does each child get?	Double-digit number divided by a single-digit number, based on a word problem.
There are 4 books in one box. The teacher has 2 such boxes. How many books are there in the two boxes altogether?	Single-digit number multiplied by a single-digit number, based on a word problem.

2.2 Predictors of learning for marginalised girls across the GEC-T portfolio

To identify the predictors of learning across the GEC-T portfolio, a series of multiple linear regressions were undertaken. The methodology and detailed results of the regressions are reported in Annex B.

The learner and teacher characteristics and behaviours included in the model are:

- primary caregiver has no education (N¹²=8,880 of N=23,826);
- household head has no education (could occasionally be the same person as the primary caregiver) (N=7,888 of N=25,090);
- girl reports that chores affect school (N= 3,051 of N=19,567);
- number of reasons travel to school is unsafe;
- age (Total N=54,749);
- overage grade status (N=21,238 of N=53,387);
- married status (N=937 or N=25,284);
- motherhood status (N=950 of N=24,882);
- reports a functional disability (N=2,802 of N=35,853);
- reports more than one functional disability (N=2,631 of N=41,779);
- number of economic barriers faced (Total N=22,488);
- home language is different than language of instruction (N=15,134 of N=21,786);
- teacher uses physical punishment as a form of discipline (N=21,292 of N=36,644);
- teacher uses detention as a form of discipline (N=3,552 of N=36,515); and
- teacher uses shouting as a form of discipline (N=7,796 of N=36,635).

The multiple regression models provide information on the additional (unique) predictive value of each of the characteristics/behaviours to learning. All the characteristics/behaviours are included in the model, and the output of the model only includes those that are statistically significant at the 5% confidence level. Each of the characteristics are not only related to learning but are also likely to be intercorrelated. Moreover, predictors of learning are not causal, and as data in the GEC-T project evaluations were not created for the purposes of analysing predictors of learning many will act as proxies for a larger construct. For example, primary caregiver education status is likely to be correlated to economic barriers at the household level.

¹² N refers to the number of cases in the database with this information available.

It is also important to note that, while the GEC-T master dataset provides a unique set of data, the data are (by design) not representative of a population group. The data are focused on marginalised girls selected according to each project's criteria. This is extremely useful for the purposes of answering questions regarding the characteristics of marginalised girls, but for the purposes of predicting learning, the variance in the data (both in terms of background characteristics and in terms of learning outcomes) will be lower than for a population-based survey. For example, marginalised girls are likely to sit in the lowest socioeconomic category in a population, and therefore the variance in their economic circumstances and in their learning outcomes will be lower than if the whole population were included in the data. For this reason, the size of specific contributions to predicting learning are likely to be different to a nationally representative group. In this way, the analysis speaks to predictors of learning for marginalised girls.

2.2.1 Mathematics

Across the global GEC-T portfolio, not speaking the language of instruction at home, the teacher using shouting as a form of discipline, and the primary caregiver having no education are negative predictors of a girl meeting the mathematics benchmark. A negative predictor indicates that generally the association across the data indicates that, as the characteristic or behaviour (independent variable) increases, learning (dependent variable) decreases.

Age and being overage for grade positively predict meeting the mathematics benchmark. This is likely to be explained by the cognitive development of older girls and the low benchmark used for the learning variable.

Figure 9: Unique contribution of predictors of meeting the mathematics benchmark (standardised beta weight coefficient)¹³



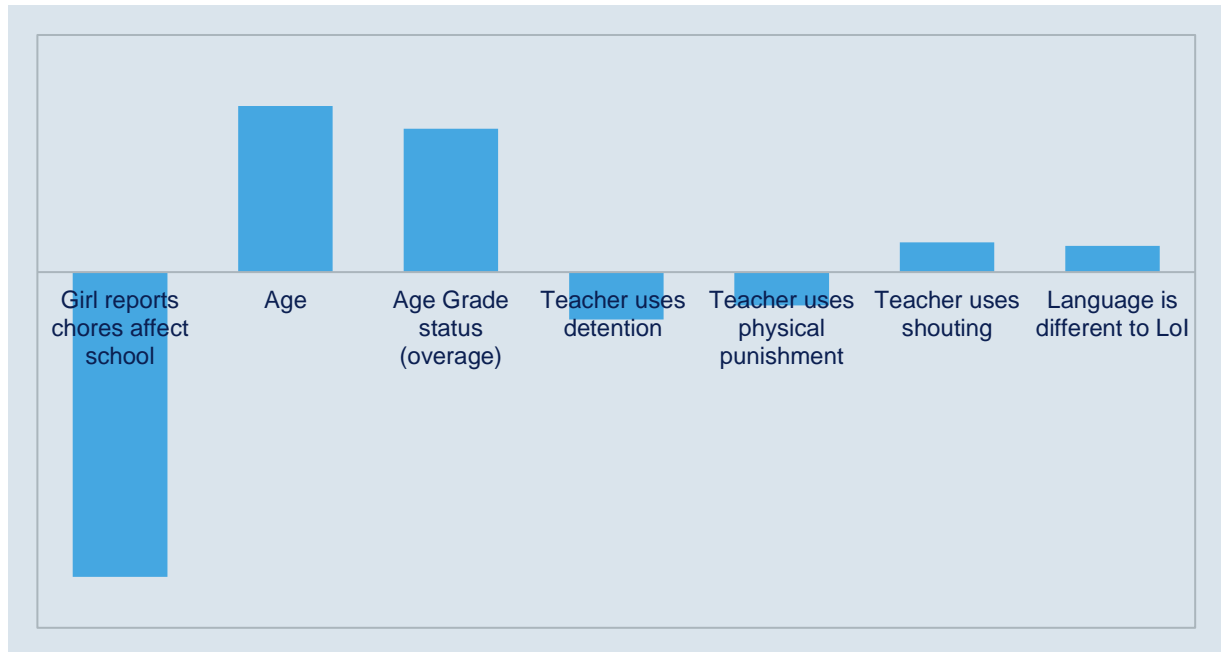
2.2.2 Reading

Across the global GEC-T portfolio, girls' reports that chores interrupt their schooling strongly negatively predict meeting the reading benchmark. As with mathematics, age and age grade status positively predict meeting the reading benchmark.

The teacher using detention and physical punishment is a negative predictor of a girl meeting the reading benchmark.

¹³ The R-squared for all predictors is 0.70. The adjusted R-squared is 0.69. Only predictors that are both statistically significant and moderate in size are reported.

Figure 10: Unique contribution to the prediction of meeting minimum reading benchmarks—GEC-T (standardised beta weight coefficient)¹⁴



2.3 How do the predictors differ by group and across contexts?

To regress all of the variables included in the model, all variables needed to be available for at least some of the sample. This was not the case in all project datasets, so the model was run where all possible variables were available across a large enough sample.

2.3.1 Age

Adolescence is a time of great change in young people's lives and introduces additional risks for many girls in low- and middle-income countries. Gender gaps widen; for many, competing demands of work and education emerge; and for some, early marriage and parenthood become lived realities (Young Lives, 2018). For this reason, the models were also run to include only girls aged 12 or above.

When running the multiple linear regressions for girls aged over 12, a reasonably clear picture emerges that, for older girls (12 years old or above), economic barriers,¹⁵ multiple functional disabilities¹⁶ being married,¹⁷ and the teacher using physical punishment as a form of discipline¹⁸ negatively predict meeting reading and mathematics benchmarks.

¹⁴ The R-squared for all predictors is 0.87. The adjusted R-squared is 0.86. Only predictors that are both statistically significant and moderate in size are reported.

¹⁵ For reading.

¹⁶ For reading.

¹⁷ For mathematics.

¹⁸ For mathematics.

2.3.2 Predictors of reading and mathematics by country

At the country level, physical, social, economic, linguistic, and teacher behavioural factors emerge as negative predictors of learning. While the country-level analysis shows a huge amount of variance in the extent to which some characteristics and behaviours predict proficiency across contexts, some constants remain. Functional disability negatively predicts learning in reading and/or mathematics across several country contexts at baseline. Facing multiple economic barriers negatively predicts learning in Tanzania and Afghanistan. Speaking a language other than the language of instruction negatively predicts reading in Afghanistan.

The teacher using physical punishment as a form of discipline negatively predicts learning, while other teacher behaviours provide mixed results.

2.3.3 Predictors of reading and mathematics by project¹⁹

The regression model used for project level data was able to use the continuous total scores (which range from 0 to 100) on reading and mathematics assessments for each project. This continuous measure of learning levels represents a more nuanced view of 'proficiency' compared to the dichotomous benchmark (where girls either meet or do not meet the benchmark). This allows more information to be fed into the regression model, and for this reason, more predictors of learning can be identified in the project-level analysis.

What emerges as important are the relationships of learning with poverty (the role of the primary caregiver and economic barriers), **gendered social norms** (the domestic burden on girls), **disability, and the use of physical punishment in schools.** These relationships remain important across all or some of the projects.²⁰

The role of the primary caregiver emerged as significant, and the primary caregiver having no education negatively predicts reading across more projects than the household head having no education. Girls reporting that chores affect school negatively predicts learning.²¹ Reporting one or more functional difficulties through the Washington Group question sets negatively predicts reading and mathematics in both projects with the highest and lowest proportions of girls with disabilities. Reporting an increased number of economic barriers at the household level negatively predicts both mathematics and reading in several projects.

For reading, physical punishment negatively predicted learning. For mathematics, shouting, physical punishment, and detention all negatively predicted learning.

¹⁹ The proportion of the variance in reading explained by the characteristics and behaviours used in the model ranged between 1% and 11% across projects, with most between 8% and 11%. The proportion of the variance in mathematics explained by the characteristics and behaviours used in the model ranged between 1% and 11% across projects, with most between 8% and 11%.

²⁰ A huge amount of variance in the extent to which other characteristics and behaviours predict reading and mathematics proficiency across projects is observable. Again, age and overage status predict learning in both positive and negative directions in different projects for reading and mathematics.

²¹ In reading.

3 Summary of the findings and implications for DFID and GEC

3.1 Summary of findings

This summary of findings moves through the evidence on learning levels globally and the predictors of learning globally, to the learning levels of GEC-T girls at baseline and the predictors of learning for GEC-T girls at the beginning of GEC Phase II.

This paper describes the relatively low number of children meeting minimum levels of reading and mathematics proficiency benchmarks in low- and middle- income countries globally, and in particular in Africa. This is supported by the GEC-T baseline data, where 44.5% of the girls did not meet grade 2 or 3 minimum proficiency levels for both mathematics and reading, despite many being in grades much higher than grade 3 and many being much older than expected for grade 3 enrolment.

The baseline data shows that marginalised girls face multiple and complex challenges. This includes limited opportunities for development in the early years, limited or delayed schooling and increased demands during young adulthood.

Limited early learning opportunities (parental caregiver with no education and less so household head has no education), **limited resources to support education** (economic barriers), **and limited exposure to language early in life that will support future learning** (home language is different than language of instruction) **all negatively predict learning globally and across the GEC-T portfolio.**

At the school level, across all of the multiple linear regression models, when the use of physical punishment as a form of discipline predict learning – it is always in a negative direction.

The largest (negative) predictor of reading outcomes at baseline is a girl reporting that chores affect school and managing the competing demands of school, work and domestic labour (chores interrupt schooling and married status) **emerges as a large negative predictor of learning for girls over 12 in the GEC-T portfolio.** This variable acts as a proxy for a range of possible interrelated social and economic factors, including access to resources within the household and gendered expectation regarding domestic labour.

Reporting one or more functional difficulties emerges across several contexts and projects as a negative predictor of both reading and mathematics. This is observable in projects with both large and small proportions of GEC-T girls reporting functional difficulties.

3.2 Implications for DFID and GEC

This report has emphasised the complexities in the relationships between predictors of learning and learning outcomes. These are relationships of association and not cause. Moreover, not all predictors of learning are amenable to education policy or programme intervention. Therefore, not all negative predictors of learning can be addressed

by an education intervention. However, schools are useful sites for intervention and can play an equalising role (Rolleston, 2014).

Based on the analysis of GEC-T baseline learning data, implications for GEC-T emerge as follows.

- 2. Projects should sensitise communities on gendered social norms and education but also provide targeted support to girls with high domestic burdens.** Sensitising communities on education and gendered social norms has been a focus of many education programmes globally, and GEC-T projects should continue work in this area. However, seeing change in gendered norms as a result of intervention is a long and slow process; the provision of support to girls who manage the competing demands of work and school should therefore be provided alongside social norm programmes.
- 3. Projects should not only focus on eliminating physical punishment in schools, but should also focus on building the capacity of teachers to develop alternative strategies.**
- 4. Who are the GEC-T Supported Girls (Outhred, Ogando Portela, Mackintosh and Keck, 2020) reported low prevalence of girls with disabilities within the GEC-T portfolio.** This was hypothesised to have happened due to two principal barriers. The first was the inability of the government schooling system to be inclusive and meet the needs of girls with disabilities, as most projects target girls through schools; and the second was the unseen nature of disability and difficulties in finding girls with disabilities due to social stigma and social norms. This paper highlights that, despite the low overall percentage of disabled girls in the GEC-T portfolio, across a number of projects with both low and high percentages of girls with disabilities, disability negatively predicts learning. This is an opportunity for DFID and GEC-T projects to ensure a clear strategy and focus towards further reaching girls with disabilities for support in learning. The success of this strategy could be analysed by conducting a similar analysis of the endline GEC-T data.

3.3 Potential further research

The findings presented in this paper suggest avenues for further research, either using the master dataset developed as part of this project, using the wider evidence base, or by collecting new qualitative and quantitative data, as follows:

1. A full benchmarking activity, qualitatively aligning the difficulty of all subtasks across projects and establishing robust comparisons across the GEC-T baseline assessments, could provide a stronger comparison on the baseline learning levels of GEC-T girls than what has been possible in this project. However, the purpose of such an exercise would need to be clear, as simply comparing learning levels across projects that are not representative of a broader population group is not useful in and of itself.

2. A review of the evidence regarding how initiatives have provided support to girls to manage the competing demands of work and school and 'what works' could provide useful information for GEC-T projects moving forward.
3. Thematic, site-specific qualitative studies on identifying the root causes of physical punishment in schools could provide information to support GEC interventions. This research could include primary qualitative data collection to better understand the conditions under which physical violence in schools takes place and strategies for eliminating violence.

List of references

- Allen, R., Elks, P., Outhred, R., and Varly, P. (2016). *Uganda's Assessment System: A Roadmap for Enhancing Assessment in Education*. Oxford: HEART.
- Baker-Henningham, H., Gardner, J., Chang, S., and Walker, S. (2009). Experiences of violence and deficits in academic achievement among urban primary school children in Jamaica. *Child Abuse & Neglect*.
- Benny, L. B. (2018). *Early is best but it's not always too late: Evidence from the Young Lives study in Ethiopia, India, Peru and Vietnam*. Oxford: Young Lives.
- Bermeo, E. (2014). South Korea's successful education system: lessons and policy implications for Peru. *Korean Social Science Reserach Council*, 135-151.
- Branigan, Amelia R, Kenneth J McCallum, and Jeremy Freese. 2013. "Variation in the Heritability of Educational Attainment: An International Meta-Analysis." *Social Forces* 109-140.
- Chen, W., Grove, W. A., and Hussey, A. (2017). *The Role of Confidence and Noncognitive Skills for Post-Baccalaureate Academic and Labor Market Outcomes*. *Journal of Economic Behavior & Organization*.
- Chen, Y., Kubzansky, L. D., & VanderWeele, T. J. (2019). Parental Warmth and Flourishing in Mid-life. *Social Science & Medicine*, 65-72.
- Claassens, L. J. (2016). The Woman of Substance and Human Flourishing. *Journal of Feminist Studies in Religion*, 5-19.
- Couch, L., and Rolleston, C. (2017, August). *Raising the floor on learning levels: Equitable improvement starts with the tail*. Retrieved from RISE programme: <https://www.riseprogramme.org/publications/raising-floor-learning-levels-equitable-improvement-starts-tail>
- Cunningham, J. W., & Mesmer, H. A. (2014). Quantitative measurement of text difficulty: What's the use? *The Elementary School Journal* , 255-269.
- Gordon B. Dahl & Lance Lochner, 2005. "The Impact of Family Income on child Achievement". *NBER Working Papers 11279*. National Bureau of Economic Research Inc.

- Durrant, J., and Ensom, R. (2012). Physical punishment of children: lessons from 20 years of research. *Canadian Medical Association Journal*.
- Gershoff, E. (2017). School Corporal Punishment in Global Perspective: Prevalence, Outcomes and Efforts at Intervention. *Psychology, health & medicine*.
- Hattie, J. A. (2009). *Visible learning: A synthesis of 800+ meta-analyses on achievement*. Oxford: Routledge.
- Heckman, J. (2006, June). Skill Formation and the Economics of Investing in Disadvantaged Children. *Science*, 312, 1900-1902.
- Heyneman, S. P. (2005). Student Background and Student Achievement: What is the right question? *American Journal of Education*, 1-9.
- Heynemann, S., & Loxley, S. (1983). The effect of primary school quality on academic achievement across twenty-nine high- and low-income countries. *American Journal of Sociology*, 88(6), 1162–1194
- Hungi, N. (2011). *Accounting for variations in the quality of primary school education*. SACMEQ.
- Johnson, J S, and E L Newport. 1989. "Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language." *Cognitive Psychology* 60-99.
- Juola, P. (n.d.). *Assessing Linguistic Complexity*. Pittsburgh: Duquesne University.
- Lenneberg, E. 1967. *Biological foundations of language*. New York: Wiley.
- Leung, A. (2016, June 17). *Field, tenor and mode - a literacy framework for all subjects*. Retrieved from <https://aliceleung.net/2016/06/17/field-tenor-and-mode-a-literacy-framework-for-all-subjects/>
- McKinsey and Company. (2017). *Drivers of Student Performance: Latin America Insights*. McKinsey .
- McWhorter, J. H. (2001). The world's simplest grammars are creole grammars. *Linguistic Typology*, 125-166.
- Nag, S. (2017). *Assessment of literacy and foundational learning in developing countries: Final report*. Oxford: Health and Education Advice and Resource Team.
- Nichols, J. (1986). Head-marking and dependent-marking grammar. *Language*, 56-117.

- OECD. (2003). How Student Performance Varies between Schools and the ROle that Socio-economic Background Plays in This. In *Learning for Tomorrow's World - First Results from PISA* (pp. 160-191). Paris: OECD.
- OECD. (2015). *What lies behind gender inequality in education?* Paris: OECD.
- OECD. (2018). *PISA for Decelopment Results in Focus*. Paris: OECD.
- Ogando Portela, M. J. (2015). Corporal punishment in schools: Longitudinal evidence from Ethiopia, India, Peru and Vietnam.
- Outhred, R and Beavis, A. (2012). Zimbabwe Early Learning Assessment Baseline Results. Australian Council for Educational Research, Melbourne.
- Outhred, R and Beavis, A. (2013). Zimbabwe Early Learning Assessment Midline Results. Australian Council for Educational Research, Melbourne.
- Outhred, R., and Lipcan, A. (2016). *Do Private Schools Produce Superior Language Outcomes?* Oxford: Oxford Policy Management.
- Outhred, R., Stubberfield, C., Ismail, I., Beavis, A., and McIntyre, J. (2013). *Social Norms and Equity in Education*. New York: UNICEF .
- Oxford Policy Management. (2019). *Impact Evaluation of EQUIP-T: Endline Quantitative Technical Report Volume 1*. Oxford: OPM.
- Pellens, T., Outhred, R., Majeed, Z., and Binci, M. (2016). *Baseline Evaluation of the Girls Education Programme*. Oxford: EDOREN.
- Penfield, W, and L Roberts. 1959. *Speech and brain mechanisms*. New York: Athenaeum.
- Rolleston, C. (2014). Learning Profiles and the 'Skills Gap' in Four Deveoping Countries: a Comparative Analysis of Schooling and Skills Development. *Oxford Review of Education*, 132-150.
- Talwar, V., Carlson, S., and Lee, K. (2011). Effects of a punitive environment on children's executive functioning: A natural experiment. *Social Development*.
- Thompson, R. S., and Nelson, C. A. (2001). Developmental Science and the Media: Early Brain Development. *American Psychologist*, 5-15.
- Tomasello, M. 2000. "Do young children have adult syntactic competence?" *Cognition* 209-253.

Williams, M, R Burder, and S Al-Baharna. 2001. "Making sense of success and failure: the role of the individual in motivation theory." In *Motivataion and second language acquisition*, by Z Dornyei and R Schmidt, 173-186. Honolulu: University of Hawaii Press.

Woldehanna, T. (2016). *Inequality, preschool education and cognitive development in Ethiopia: Implication for public investment in pre-primary education*. International Journal of Behavioural Development.

World Bank. (2019, January 22). The Education Crisis: Being in school in not the same as learning.

Young Lives. (2018). *Snapshot: Adolescence provides a second crucial window of opportunity to address child poverty*. Oxford: Young Lives.

Annex A Learning levels and predictors of learning globally

This section of the paper reviews global learning levels as reported by UIS. The section then reviews the global literature on predictors of learning across contexts. To establish the predictors of learning, learning levels need to be measured through an appropriate assessment for each learner, and data on background and school characteristics need to be collected and linked to individual learners. Learner-level linked data is rarely available in low- and middle-income countries and is expensive to collect. In many of the GEC-T countries, such data are not available. In FCAS environments, the evidence gap is even greater, as access to a representative sample of schools is often not possible. Therefore, this section of the paper reviews the global literature on predictors of learning across a variety of contexts.

There is a growing concern that, across many low- and middle-income countries, many children are not acquiring the necessary skills expected of them. In Kenya, Tanzania, and Uganda, three-quarters of third graders cannot read and understand a simple sentence. In rural India, nearly three-quarters of third graders cannot solve a two-digit subtraction problem (World Bank, 2019).

Variations in learning proficiency are observed across countries, and the predictors of learning and the size and direction of gaps in learning differ by social group across contexts. This is because perceived value, interest, and attitudes towards education are also shaped by the nature of interaction with significant others, the learning environment, and the broader external context (Williams *et al.*, 2001). Both the real and perceived benefits of schooling are affected by gender, age, and wealth (Rolleston, 2014, p. 135). The extent to which individuals and schools can convert endowments into learning (and convert learning into academic and workforce outcomes) differs by individual background factors (Outhred & Lipcan, 2016).

Formal schooling can play an equalising role, especially for the most disadvantaged, as the schooling environment often represents the largest influence on children outside their households (Rolleston, 2014, p. 135). However, many low- and middle-income countries have struggled to maintain quality while expanding access to education, and this has had the greatest impact on the poor (Rolleston, 2014, p. 135).

It is within this context that we review the learning levels globally.

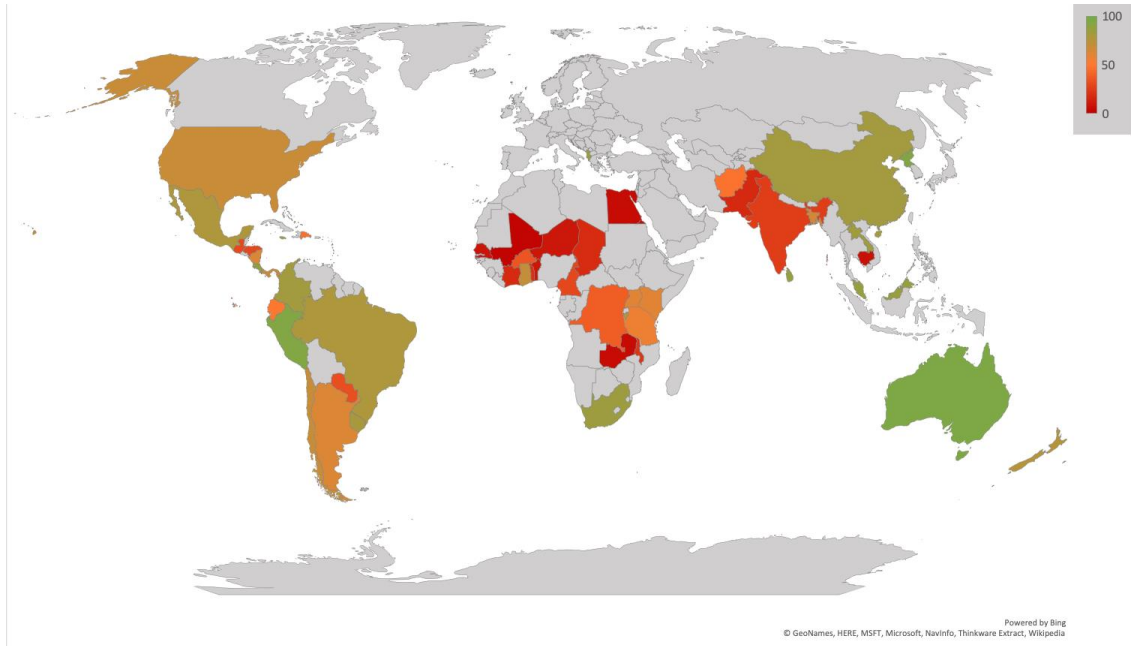
A.1 Global learning levels

To review global learning levels, UIS and PISA-D data were downloaded (see methodology). These data were placed into a geographical mapping tool with colour coding to represent countries with high (green), medium (orange), and low (red) percentages of learners who meet benchmarks.

As can be seen in Figure 11 and Figure 12, **regional patterns of proficiency are observable for both reading and mathematics**, with extremely low percentages of

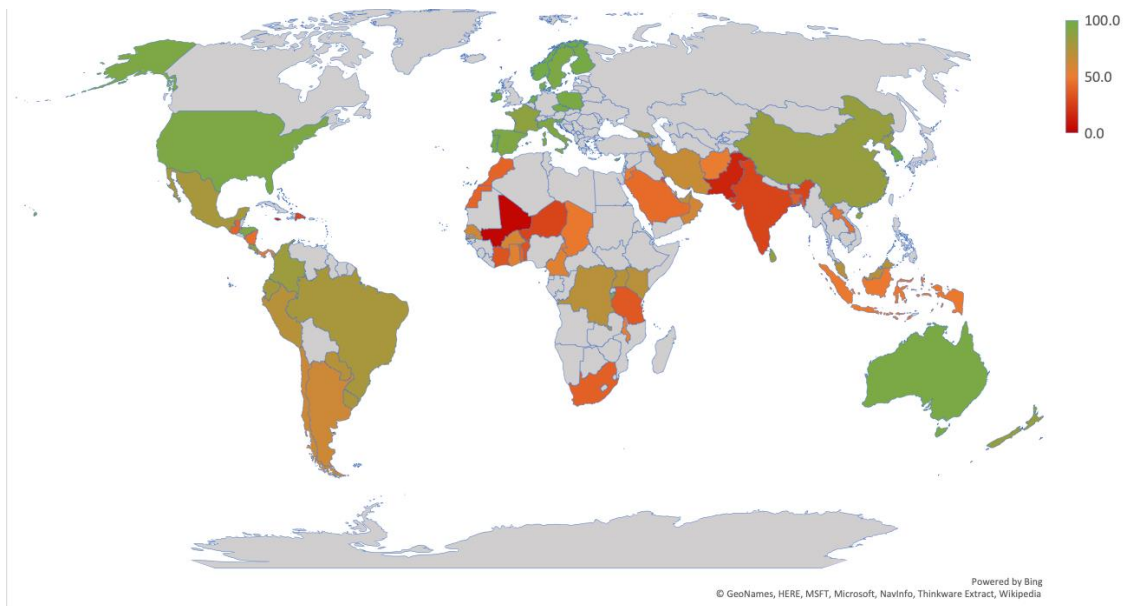
children in Eastern and Central Africa, Western Africa, and South Asia meeting the reading benchmarks. In mathematics, extremely low percentages of children across Africa and South Asia meet the benchmarks. Reasonably low levels of children meet the benchmarks in reading in the Americas and in mathematics in South America. Where data are available, higher rates of proficiency are observed in Europe, the United States, Australia, and New Zealand.

Figure 11: Percentage of children who meet reading benchmarks globally



Source: UIS and PISA-D data

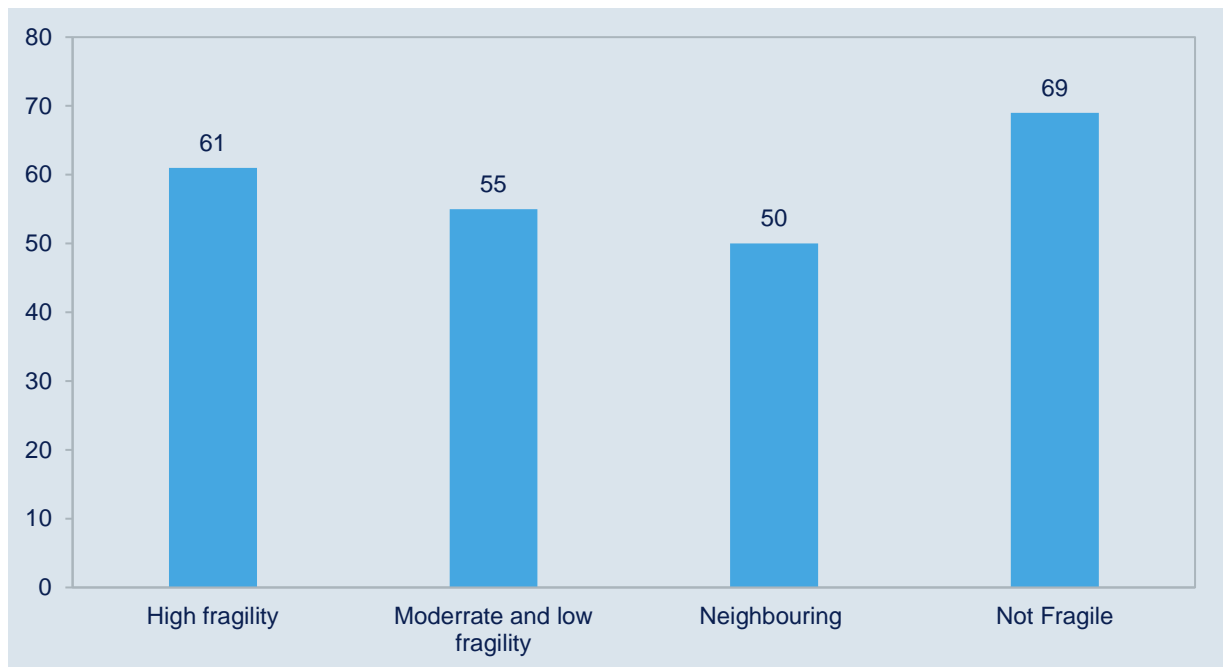
Figure 12: Percentage of children who meet mathematics benchmarks globally



Source: UIS and PISA-D data

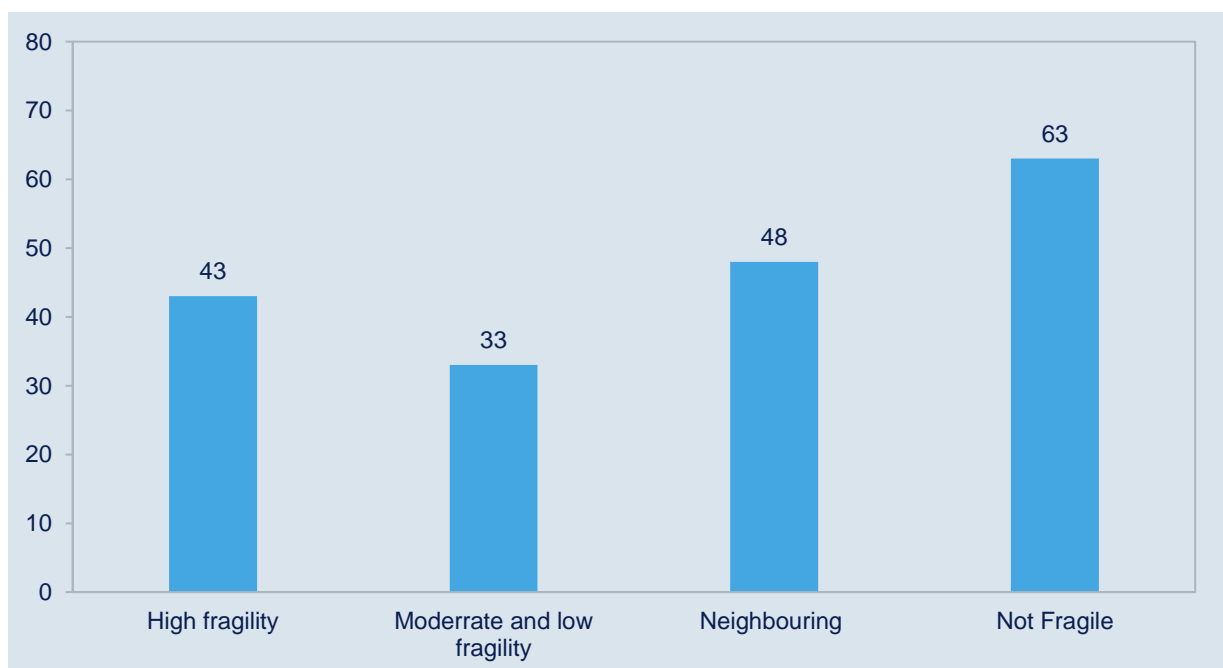
The percentage of children who meet benchmarks in reading and mathematics is lower in fragile states and neighbouring states compared to non-fragile states (Figure 13 and Figure 14).

Figure 13: Average percentage of children achieving minimum proficiency in mathematics, by fragility status



Source: UIS data

Figure 14: Average percentage of children achieving minimum proficiency in reading, by fragility status



Source: UIS data

A.2 Predictors of learning by context²²

In this section of the report, we discuss the evidence base regarding the role of schooling, socioeconomic status (SES), gender, age, and linguistic factors in predicting learning outcomes. However, these factors are interrelated and interdependent and cannot be viewed in isolation.

Heckman states that cognitive development and skill formation are influenced by '*an interaction between genetics and individual experience*' (2006). He goes on to explain that the underlying neural pathways required for skill mastery follow hierarchical rules, and later attainments are built on foundations that are laid down early in life (Thompson & Nelson, 2001) Cognitive, linguistic, social, and emotional competencies are interdependent, shaped powerfully by experience; while humans are able to adapt throughout life, neural circuits are set early on and are most open to environmental influence in the early years (Heckman, 2006).

The role of schooling in supporting the development of knowledge and skills is well established. Formal schooling plays a role in the development of cognitive skills, from which productive abilities are derived (Rolleston, 2014, p. 135).

Schools can reinforce inequalities or they can have an equalising effect. For example, Young Lives research shows that some children who grow up in a poor environment reach high levels of learning and '*schools play a big role in explaining these differences*' (Woldehanna, 2016). Studies from high-income countries clearly demonstrate that the provision of early childhood education supports later development and reduces economic inequality (Woldehanna, 2016). The body of evidence strongly supports the case that cognitive and non-cognitive stimulation in early life through early childhood education are critical for long-term skills development (Woldehanna, 2016). Primary and lower secondary schooling plays a particularly important role in the development of basic cognitive skills, which become the foundation for future and more complex cognitive skills, including problem solving and communication. The development of these skills through formal schooling plays a role in improving young people's prospects of labour market participation (King and Palmer 2006; Rolleston, 2014, p. 135).

While the home environment remains the strongest predictor of learning, the school environment typically represents the second largest influence on children outside of their homes (Rolleston, 2014, p. 135).

An analysis of SACMEQ data finds that pupils who never repeated grades were likely to achieve better than pupils who repeated grades across all 15 SACMEQ school systems in reading and in 13 of the 15 SACMEQ school systems in mathematics (with the exception of Malawi and Swaziland) (Hungu, 2011). Days absent from school and homework completion were also significant predictors.

²² While UIS and PISA-D data provide a useful global, regional, and country context profile of the proportions of children who meet minimum benchmarks, most of these data are not linked to individual, school, or contextual data. Therefore, what predicts learning cannot be explained by the data. In this section of the report, we summarise the global literature on predictors of learning across contexts using evidence from across assessment programmes and school surveys.

SES matters across contexts, but the size of the explanatory power differs by context.

There are ongoing debates regarding the role of schooling versus household income in predicting performance, but overall research suggests that the explanatory power of school quality differs by context. Heyneman and Loxley (1983) find that school quality was a more important indicator of achievement in poorer countries (Heyneman, 2005).

SES predicts learning across SACMEQ countries. A multivariate analysis of the predictors of learning across SACMEQ participating countries found that pupil SES significantly predicts learning in both reading and mathematics. Pupils from richer families achieved better score than pupil from poor families across all of the 15 school systems for reading and across 13 for mathematics. The largest differences in reading scores between rich and poor pupils were found in South Africa, Mauritius, and Zimbabwe (Hungu, 2011). For reading, meals per week, household tasks, number of siblings, and having books at home also predicted learning. In Nigeria, two studies (one in Lagos and one in northern Nigeria) found that SES predicted learning outcomes (Outhred and Lipcan, 2016 and Pellens, Outhred, Majeed, and Binci, 2016).

SES is not only about wealth. SES encompasses a set of social and economic capital bestowed upon an individual, usually through the status of one's birth. This capital includes non-wealth endowments such as parental education, parental motivation, early learning opportunities, language exposure, family formation, and access to reading materials.

Social and economic factors interact. For example, parental education is a cause and consequence of wealth, and parental education and motivation interact with household and early learning opportunities. Studies indicate that parental warmth promotes offspring functioning across multiple domains of well-being in mid-life (Chen et al, 2019) and the extent to which children are exposed to language from parents and caregivers from a young age predicts development (Tomasello, 2000).

The role of gender in predicting learning differs across contexts and across the learning continuum. Gender interacts with class and other social factors across contexts. Over the past century, gender gaps in education and employment, educational attainment, pay, and labour market participation have significantly narrowed in OECD countries (OECD, 2015). The OECD report on what lies behind gender inequality states: *'This one fact implies another: that aptitude knows no gender. Given equal opportunities, boys and girls, men and women have equal chances of achieving at the highest levels.'*

Across Africa gender as a predictor of learning differs in the size and direction of the gap. In Tanzania, the impact evaluation of DFID's Education Quality Improvement Project found that teacher interactions with pupils became more gender-balanced over the project period, and at the end of the evaluation period an attainment gap opened up in Kiswahili (with girls performing better than boys), while the boys' lead in mathematics narrowed (Oxford Policy Management, 2019).

Evidence from northern Nigeria finds that gender is correlated with learning. Boys were found to perform better than their female counterparts, although the size of the correlation was small after controlling for school-level characteristics (Pellens, Outhred, Majeed, & Binci, 2016). These relationships are akin to gendered norms, whereby attitudes towards girls and the expectations of girls and boys differ (Pellens, Outhred, Majeed, & Binci, 2016).

In southern Nigeria (Lagos state), gender emerged as a particularly important predictor of language outcomes in public schools, with boys performing significantly worse than girls (Outhred & Lipcan, *Do Private Schools Produce Superior Language Outcomes?*, 2016).

In SACMEQ data, gender was also significant in predicting performance in Kenya, Malawi, Mozambique, Tanzania, Uganda, and Zambia, with boys outperforming girls in reading and mathematics. However, in the Seychelles, girls outperformed boys in the two subjects. In Mauritius, Namibia, and Zanzibar, girls did better than boys in reading, while boys did better than girls in mathematics.

The role of age is important and not straightforward. Age is a nuanced characteristic; while cognitive development takes place over time and the opportunity to improve proficiencies are biologically present, a lack of opportunity to develop foundational skills in the early years can place older children behind in their learning. This is particularly important in contexts where the language of instruction is different to the home language of the learner. There is a strong evidence base supporting the notion that maturational constraints impede the completeness of language development in learners after puberty (Lenneberg 1967; Penfield and Robertson Hahne and Friderici 2001; Johnson and Newport 1989).

This is evidenced in a public/private school comparison study in Lagos, where more overage children who do not speak the language of instruction in the home exist in public schools. More disadvantaged children tend to be older than others in the same grade, and therefore age is often confounded with SES. This was reflected in the study's findings that age is positively correlated with performance in private schools and negatively correlated with performance in public schools (Outhred & Lipcan, *Do Private Schools Produce Superior Language Outcomes?*, 2016).

In northern Nigeria, learning levels are extremely low, and the vast majority of learners in grade 2 were found to be overage. Older pupils performed better than younger pupils, although even the highest performing pupils were not reaching grade 2 level reading expectations in Hausa or English (Pellens, Outhred, Majeed, & Binci, 2016).

The language spoken at home is a complex and important predictor of learning. In Tanzania, pupils from households who speak Kiswahili at home perform far better than those who do not in both Kiswahili and mathematics, even after taking poverty into account. Across all 15 SACMEQ participating schooling systems, speaking the language of instruction emerged as a significant predictor of learning in reading and mathematics (Hungu, 2011). Pupils who spoke the language of instruction at home were more often estimated to achieve better than pupils who rarely or never spoke the language of instruction in the home across systems (Hungu, 2011). Speaking the language of instruction at home also emerged as a significant predictor of learning in reading and mathematics in Tanzania (Oxford Policy Management, 2019).

Annex B Multiple linear regression methodology and results

Based on the data available from project level evaluations, individual and teacher characteristics and behaviours were selected for the multiple linear regression.

The multiple linear regression decomposes the unique contribution of each variable as a predictor of learning outcomes. Using the stepwise approach with forward selection, only statistically significant predictors of learning are included in the model. The direction and size of the correlation, the statistical significance, and the extent to which the correlation explains the variance in the data are then reported on.

In the stepwise multiple regression, Model 1 is run and reports on the biggest predictor. Model 2 then adds the next biggest predictor, and a new model is run until no statistically significant relationships remain.

Some of the teacher- and school-level variables were only available for projects that have teaching quality or school governance and management as intermediate outcomes. We assessed this and used a limited set of variables to be included in any regression analyses (Table 4).

Table 4: Individual-, teacher-, and school-level factors to be used in the analysis

Individual-, household-, and community-level factors	Teacher-level factors
Age	Teacher uses physical punishment
Girl reports that chores affect school	Teacher uses detention
Primary caregiver has no education	Teacher uses shouting
Household head has no education	
Single disability status	
Multiple disability status	
Age grade status	
Married status	
Motherhood status	
Language status (Home language is different from the language of instruction)	
Number of economic barriers faced in the household	

Table 5: Positive and negative relationships between individual and teacher characteristics and meeting the reading benchmark, by country

Characteristics/behaviours included in the model	Afghanistan	Kenya	Mozambique	Sierra Leone	Somalia	Tanzania	Zimbabwe
Primary caregiver has no education	-		-				
Household head has no education							
Girl reports that chores affect school		-					
Number of reasons travel to school is unsafe	+ ²³	+ ²⁴					-
Age	+	-		+		-	+
Overage grade status			+	+		+	+
Married							
Mother							
Reports a functional disability							
Reports more than one functional disability				-		-	
Number of economic barriers faced						-	
Home language is different from the language of instruction	-						
Teacher uses physical punishment as a form of discipline							
Teacher uses detention as a form of discipline							-
Teacher uses shouting as a form of discipline							

Table 6: Positive and negative relationships between individual and teacher characteristics and meeting the mathematics benchmark, by country

Characteristics/behaviours included in the model	Afghanistan	Kenya	Mozambique	Sierra Leone	Somalia	Zimbabwe
Primary caregiver has no education						
Household head has no education						
Girl reports that chores affect school						
Number of reasons travel to school is unsafe						
Age	+	-	-	-	-	

²³ This is unexpected.²⁴ This is unexpected.

Characteristics/behaviours included in the model	Afghanistan	Kenya	Mozambique	Sierra Leone	Somalia	Zimbabwe
Overage grade status		-	-	-		+
Married	-					
Mother						
Reports a functional disability				+ ²⁵		
Reports more than one functional disability						-
Number of economic barriers faced	-					
Home language is different from language of instruction	-					
Teacher uses physical punishment as a form of discipline	-					-
Teacher uses detention as a form of discipline		+	-			
Teacher uses shouting as a form of discipline			+			

Table 7: Positive and negative relationships between individual and teacher characteristics and overall achievement in EGRA/SEGRA, by project

Characteristics/behaviours included in the model	AKF	BRAC	CAMFED – Multi-Country	CAMFED – MC (Zimbabwe)	CAMFED Tanzania	CARE	Plan	Relief	World Vision
Primary caregiver has no education	-				-	-		-	
Household head has no education	-								
Girl reports that chores affect school	-	-							
Number of reasons travel to school is unsafe						-		+	
Age	+	+	-	-	-				+
Overage grade status	+						-	+	+
Married									
Mother									
Reports a functional difficulty	-		-						
Reports more than one functional difficulty		-		-	-				

²⁵ This is unexpected.

Characteristics/behaviours included in the model	AKF	BRAC	CAMFED – Multi-Country	CAMFED – MC (Zimbabwe)	CAMFED Tanzania	CARE	Plan	Relief	World Vision
Number of economic barriers faced	-		-						
Home language is different than language of instruction		-							
Teacher uses physical punishment as a form of discipline		-							
Teacher uses detention as a form of discipline	+		-	-					
Teacher uses shouting as a form of discipline	-	-	+						

Table 8: Positive and negative relationships between individual and teacher characteristics and overall achievement in EGMA/SeGMA, by project

Characteristics/behaviours included in the model	AKF	BRAC	CAMFED – MC (Tanzania)	CAMFED – MC (Zimbabwe)	CAMFED Tanzania	CARE	Plan	Save Mozambique	World Vision
Primary caregiver has no education								-	
Household head has no education	-								
Girl reports that chores affect school									
Number of reasons travel to school is unsafe	+	-							
Age	+	+	-	-	-				-
Overage grade status	+				-	-		+	+
Married	-								
Mother									
Reports a functional difficulty				-	-				
Reports more than one functional difficulty		-	-				-		-
Number of economic barriers faced	-				-				
Home language is different from language of instruction		-				+			
Teacher uses physical punishment as a form of discipline		-		-		-			
Teacher uses detention as a form of discipline			-						

Characteristics/behaviours included in the model	AKF	BRAC	CAMFED – MC (Tanzania)	CAMFED – MC (Zimbabwe)	CAMFED Tanzania	CARE	Plan	Save Mozambique	World Vision
Teacher uses shouting as a form of discipline	-	-							-

Annex C Projects included in proficiency benchmarks

Table 9 and Table 10 show the projects that were included in the proficiency benchmarks for reading and mathematics. The tables show the percentage of girls who reach the reading and mathematics benchmark in each project.

Table 9: Percentage of girls reaching the mathematics benchmark, by project

Project	Percentage of girls reaching mathematics benchmark
1 AKF	29.6
2 Avanti	72.7
3 BRAC	56.1
5 CAMFED – MC (Zambia)	37.3
8 CARE	90.2
11 DLA Ghana	51.7
12 DLA Kenya	54.3
13 DLA Nigeria	42.1
14 EDT	21.6
15 HPA	56.5
17 LCD	29.9
20 Mercy Corps Nigeria	74.8
23 Plan	8.8
24 Relief	96.0
25 Save DRC	47.5
26 Save Mozambique	14.9
28 Viva	74.0
30 World Vision	64.5

Table 10: Percentage of girls reaching the reading benchmark, by project

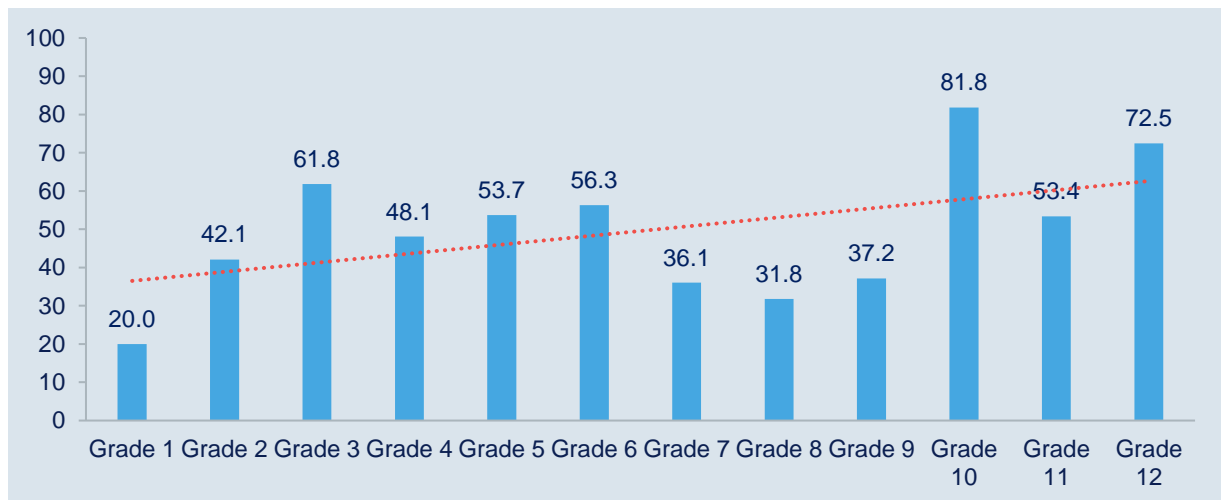
Project	Percentage of girls reaching reading benchmark
1 AKF	14.0
2 Avanti	22.4
3 BRAC	23.6
4 CAMFED – MC (Tanzania)	69.9
5 CAMFED – MC (Zambia)	23.2
6 CAMFED – MC (Zimbabwe)	49.1
7 CAMFED Tanzania	62.2

Project	Percentage of girls reaching reading benchmark
9 Childhope	50.5
11 DLA Ghana	39.2
12 DLA Kenya	80.7
13 DLA Nigeria	6.3
14 EDT	8.6
15 HPA	58.7
17 LCD	39.4
18 LINK	23.9
19 Mercy Corps NE	28.9
22 PEAS	35.5
23 Plan	55.2
24 Relief	82.3
25 Save DRC	10.1
26 Save Mozambique	15.6
27 Varkey	16.5
28 Viva	57.0
29 VSO	59.3
30 World Vision	12.6
31 WUSC	29.3

Annex D Learning levels by school and individual factors

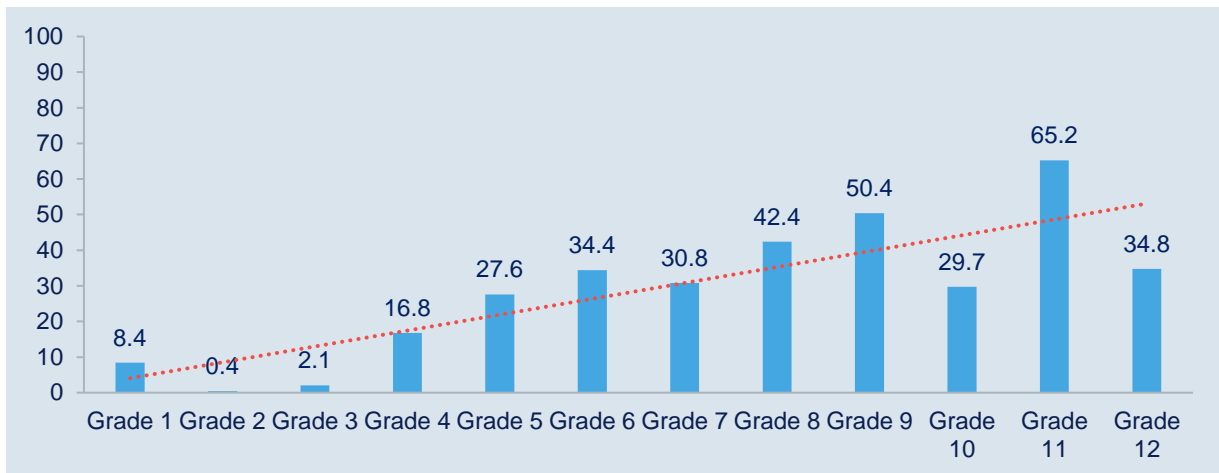
Indicators of school factors available in the GEC-T baseline master dataset include grade of enrolment and overage status.²⁶ As is to be expected, it is generally the trend that, as the grade of enrolment increases, the percentage of girls who meet the benchmark also increases (indicated by the trend line in Figure 15 and Figure 16).

Figure 15: Percentage of girls who meet the mathematics benchmark, by grade



²⁶ These are both individual and school factors.

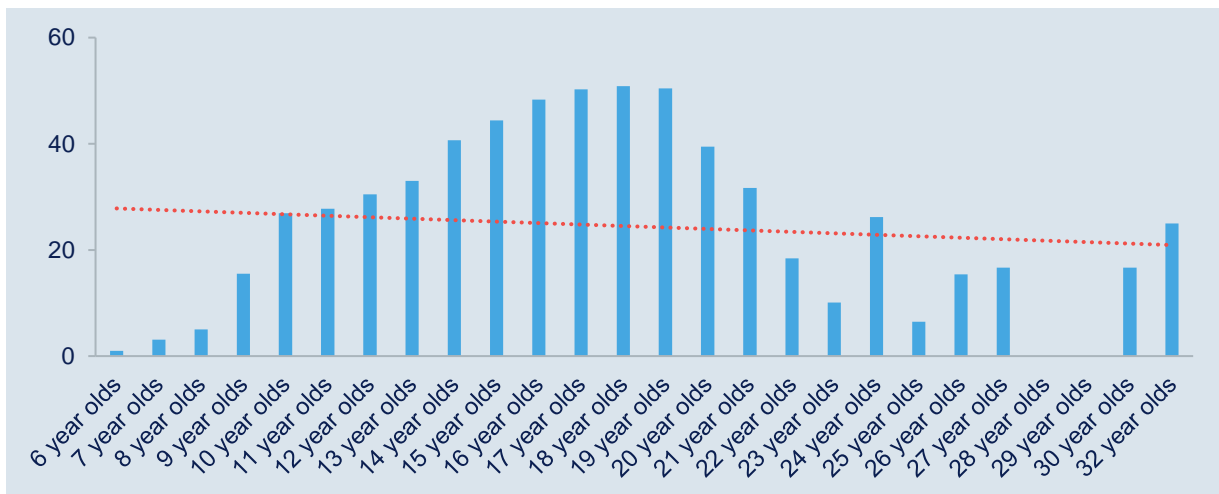
Figure 16: Percentage of girls who meet the reading benchmark, by grade



The same trend is not observable between the percentage of girls meeting reading and mathematics benchmarks and age.

For reading, the percentage of girls meeting the benchmark increases with age until the age of 19. The percentage of girls meeting the reading benchmark dramatically decreases for girls who are over 20 years old.

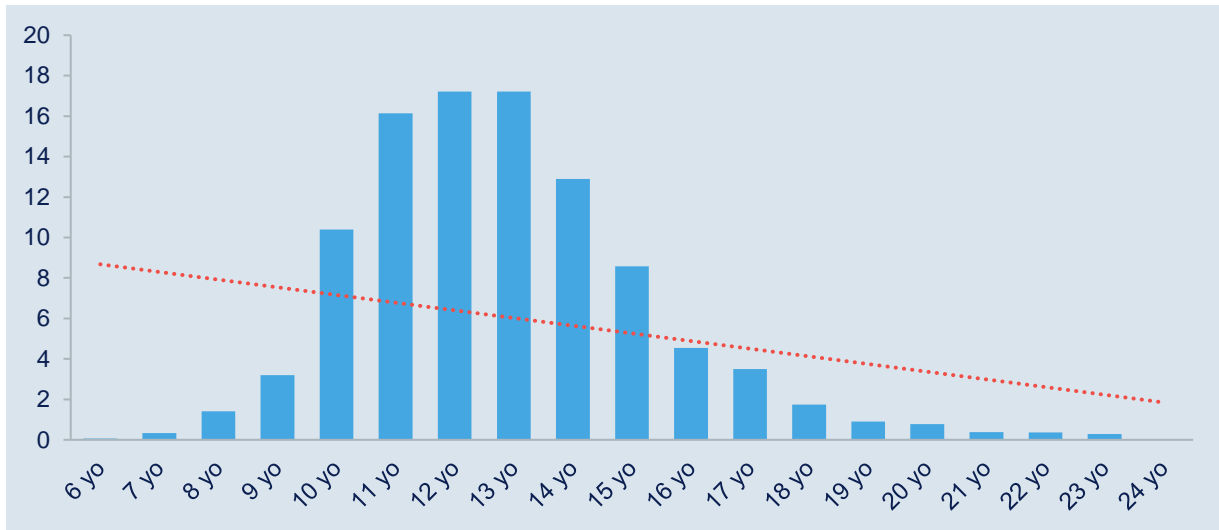
Figure 17: Percentage of girls who meet the reading benchmark, by age



The data show very low learning levels for older children for mathematics. After the age of 13, learners seem to make less and less progress.²⁷ For mathematics, the percentage of girls meeting the benchmark increases with age until the percentage peaks at 12 and 13 years of age. The percentage of girls meeting the mathematics benchmark is extremely low after 17 years of age.

²⁷ The GEC-T data is not longitudinal, and therefore differences by age represent a sample of girls' learning at one point in time.

Figure 18: Percentage of girls who meet the mathematics benchmark, by age



The percentage of girls meeting the reading benchmark is higher for girls who are enrolled in grades higher than expected for their age. The percentage of girls meeting the benchmark is lower for correct age/grade enrolment, lower still for overage, and the lowest for severely overage.

For mathematics, the percentage of girls meeting the benchmark is about the same across age/grade status groups.

Figure 19: Percentage of girls who meet the reading benchmark, by age/grade status

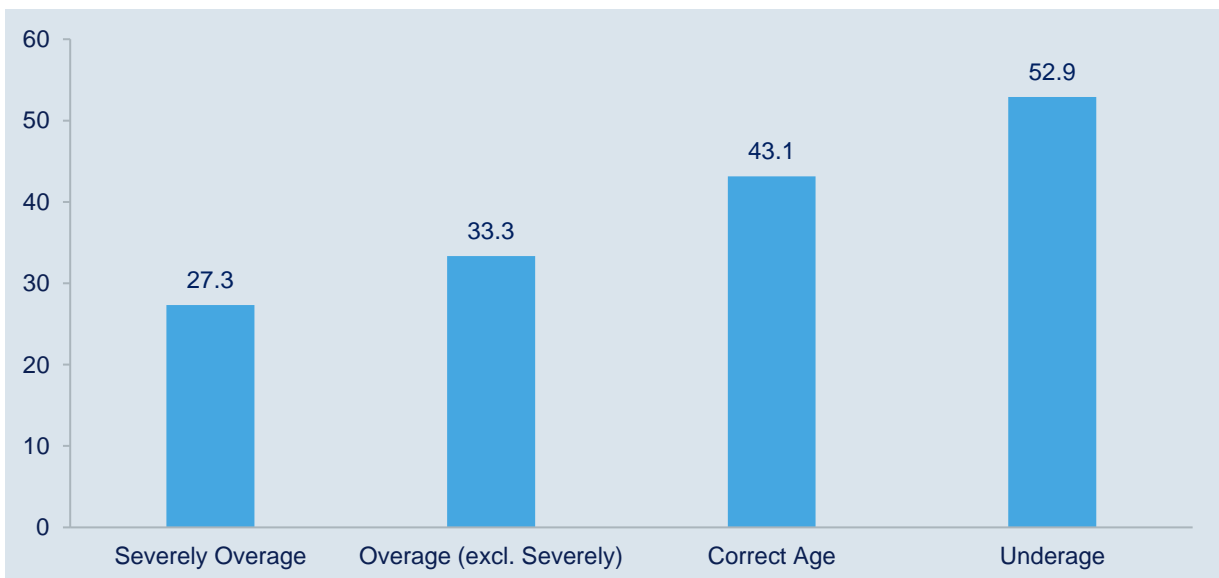
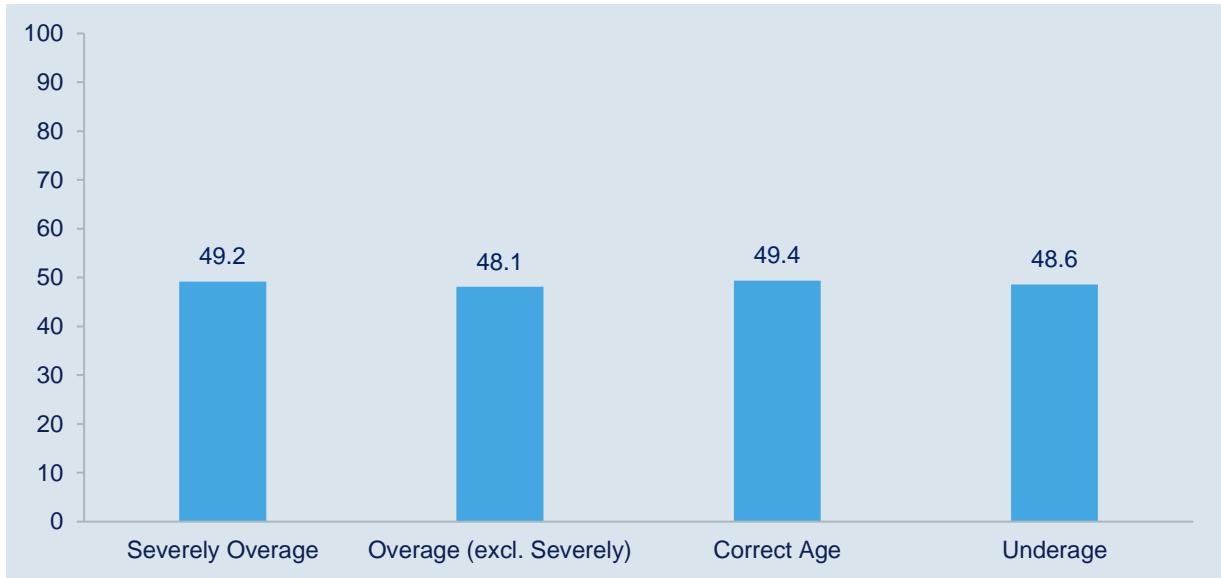
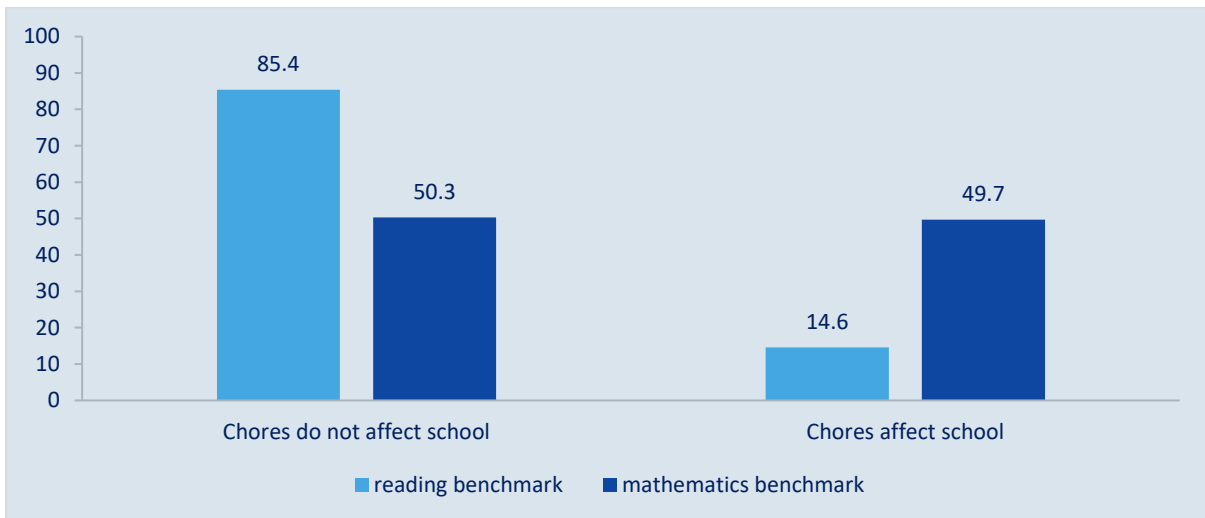


Figure 20: Percentage of girls achieving mathematics benchmark, by age/grade status



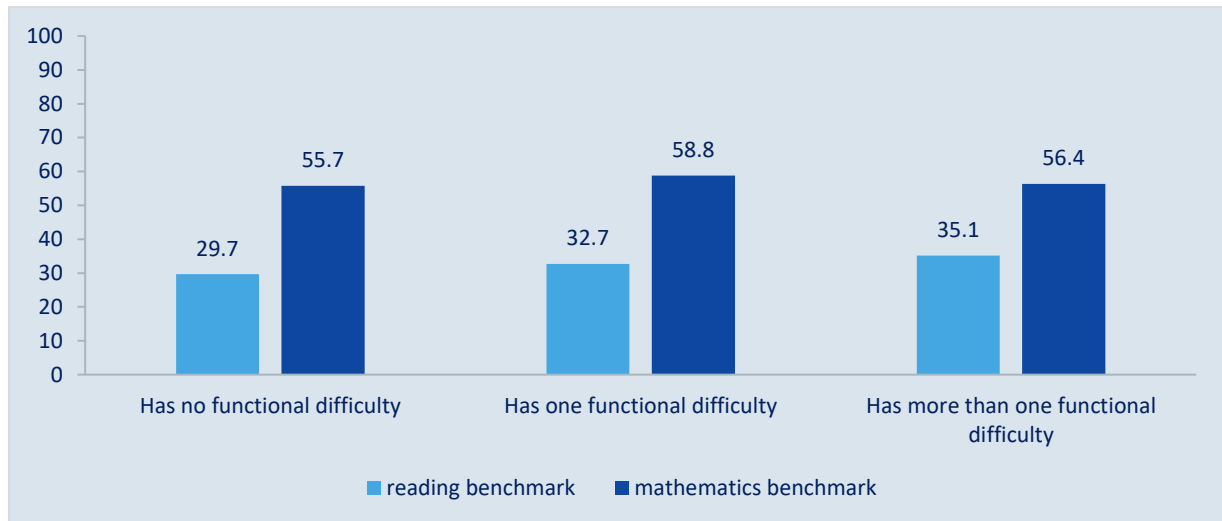
Girls were asked if completing chores affects their schooling. In mathematics, around the same percentage of girls are meeting the benchmark, whether they state that chores affect school or not. For reading, girls who report that chores affect school are much less likely to meet the benchmark than those who report that chores do not affect their schooling.

Figure 21: Percentage of girls who meet reading and mathematics benchmarks, by whether chores affect their schooling



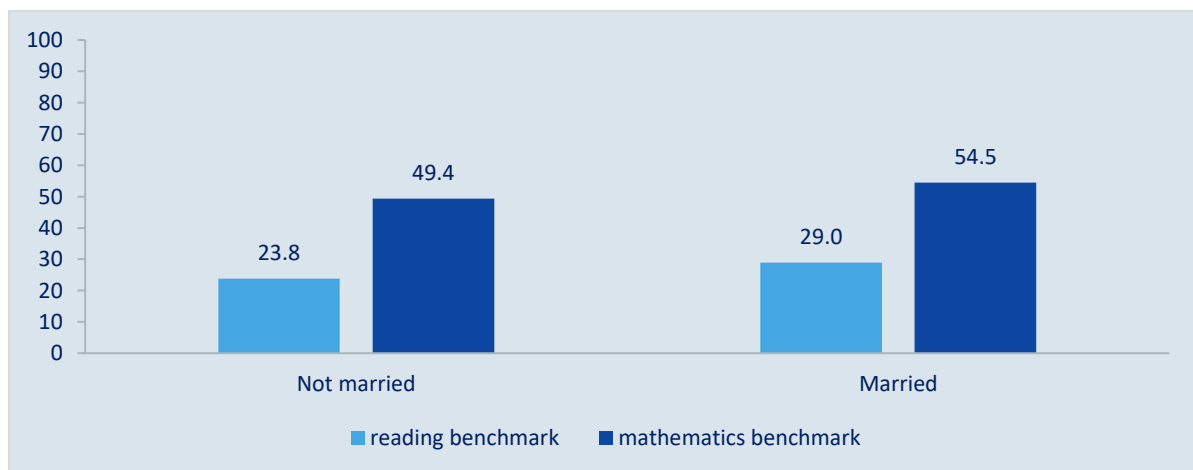
In both mathematics and reading, girls who report having a functional difficulty have a similar likelihood of meeting the benchmark compared to girls who report not having any functional difficulties.

Figure 22: Percentage of girls who meet reading and mathematics benchmark, by functional difficulty



The percentage of married girls who meet the reading and mathematics benchmarks is higher compared to non-married girls. Similar patterns are observable for motherhood status, likely because girls who are married and girls who are mothers are older compared to other girls.

Figure 23: Percentage of girls who meet reading and mathematics benchmarks, by marital status



The percentage of girls who meet the reading benchmark decreases as the number of economic barriers faced by the girl's household increases. The percentage of girls who meet the mathematics benchmark does not change as much as the number of economic barriers faced by the girl's household increases.

Figure 24: Percentage of girls who meet reading and mathematics benchmarks, by number of economic barriers