Working Paper 152

Young Lives

Learning Outcomes of Children Aged 12 in Ethiopia:

A Comparison of Two Cohorts

Tassew Woldehanna and Aregawi Gebremedhin

www.younglives.org.uk

MAY 2016

Learning Outcomes of Children Aged 12 in Ethiopia: A Comparison of Two Cohorts

Tassew Woldehanna and Aregawi Gebremedhin

ISBN 978-1-909403-67-3

A catalogue record for this publication is available from the British Library.

Funded by

UKaid
from the British people

Irish Aid
An Roinn Gnóthaí Eachtracha agus Trádála
Department of Foreign Affairs and Trade

Young Lives, Oxford Department of International Development (ODID), University of Oxford, Queen Elizabeth House, 3 Mansfield Road, Oxford OX1 3TB, UK

Tel: +44 (0)1865 281751 • E-mail: younglives@younglives.org.uk

Contents

Abstract

The paper examines inequality of learning outcomes among children of the same age (age 12) but seven years apart – in 2006 and 2013. It employs statistical analysis and uses data from the Young Lives Older Cohort who were aged 12 in 2006 (Round 2 of the Young Lives survey) and the Younger Cohort who were aged 12 in 2013 (Round 4), focusing on the Peabody Picture Vocabulary Test (PPVT), mathematics and common reading test results. Changes in test scores for the PPVT, mathematics and reading items were disaggregated by gender, location, region, caregiver education, and whether children attended private or government school. We calculated which groups were showing declining or improving levels of learning more and tried to test the significance statistically. We found that the reading and maths competencies of children had fallen substantially; and girls experienced a higher rate of decline in levels of learning in maths and reading as compared to boys, as did children living in sites in SNNPR and Oromia (compared to children from other regions), rural children (compared to urban children), children in government schools (compared to non-government schools), and children whose caregivers had received little or no education. While the average PPVT raw score of children had increased, rural children and children from SNNPR had seen a decline in their scores; and increased inequality in test results between children in private and government schools and rural and urban children have also been observed. There is an urgent need to halt the increasing learning inequality among children, which may result in inequality in adulthood. The paper proposes further research to identify the in-school and out-of-school factors that contribute to declining levels of learning and widening inequality in learning outcomes.

The authors

Tassew Woldehanna is Associate Professor, Department of Economics, Addis Ababa University and Principal Investigator of Young Lives, Ethiopia, Ethiopian Development Research Institute (EDRI).

Aregawi Gebremedhin is Assistant Researcher, Young Lives, Ethiopia, Ethiopian Development Research Institute (EDRI).

Acknowledgements

We wish to thank the external reviewers for helpful comments and Isabel Tucker for copy-editing the report. However, responsibility for any errors is ours.

About Young Lives

Young Lives is an international study of childhood poverty, following the lives of 12,000 children in 4 countries (Ethiopia, India, Peru and Vietnam) over 15 years. **www.younglives.org.uk**

Young Lives is funded by UK aid from the Department for International Development (DFID) and co-funded by Irish Aid from 2014–16. This work was funded by the Children's Investment Fund Foundation as part of a project supporting the use of research to develop effective adolescent sexual and reproductive healthcare policies.

Introduction

Education is a basic and significant factor in the development of children, communities and countries. It will help break the intergenerational chains of poverty because it is intrinsically linked to all development goals, including empowering women, improving child and maternal health, reducing hunger, fighting the spread of HIV and diseases of poverty, spurring economic growth and building peace. Owing to the fact that the educational status of individuals is measured partly by the time they spend in school, many countries and development partners have been focusing on enabling children to attend school, and they have made concerted efforts to help the enrolled students continue at school. However, until recently relatively little was done in the Ethiopian context using longitudinal data to assess changes in learning outcomes at school.

Measuring the educational status of children through school enrolment and completion rates per se may lead to unintended policy outcomes. A country could achieve a primary school enrolment rate of 100 per cent but its children might not have acquired the necessary literacy and numeracy skills after completing their primary schooling. According to findings from International Educational Assessment's *Trends in International Mathematics and Science Study at the Fourth and Eighth Grades*, the performance of children from many of the low-income nations was by far below the mean for countries belonging to the Organization for Economic Cooperation and Development (OECD) (Mullis et al. 2004). Moreover, according to the estimations of the 2013/14 Education For All Global Monitoring Report, some 250 million primary-school-age children around the world, 130 million of whom have spent at least four years in school, were not able to meet the minimum learning standards in reading and mathematics (UNESCO 2015). This figure could have been much worse had it had been measured more exactly; many countries do not measure basic reading or arithmetic in primary grades.

The Government of Ethiopia has a vision to transform the country into a middle-income economy by the year 2025, which demands transformation of the economy through the application of science and technology as instruments to create wealth. This calls for continued expansion of and equitable access to high-quality general education (Grades 1 to 12) with promising foundations in science and mathematics. In phase I of the Growth and Transformation Plan I (GTP-I, 2010/11–2014/15), one of the Government's priorities was to improve education and ensure its quality and efficiency at all levels. To realise this aim, the General Education Quality Improvement Package (GEQIP) was launched in 2008, and was due to have been fully implemented by 2015. The Ministry of Education (MOE), as part of the fourth stage of its Education Sector Development Programme (ESDP IV), implemented from 2010/11 to 2014/15, has developed a strategy of improving science and mathematics education particularly applicable to general education. Its goal is to enhance the quality of science and mathematics education and improve the performance of students, which will help them acquire the necessary skills, knowledge, awareness and problem-solving competencies in sciences and mathematics when they are streamed into either Technical and Vocational Education and Training (TVET) or the fields of engineering, technology, natural and computational science studies in higher education (MOE 2010).

In Ethiopia, access to primary education has increased significantly in recent years; for instance the numbers of schools have been increasing, with an average annual growth rate of 5 per cent since 2008/9. Similarly, the net enrolment rate for the first cycle of primary

education (Grades 1 to 4) has increased from 83 per cent in 2008/9 to 95.5 per cent in 2012/3 and from 83 per cent to 85.9 per cent for the whole of primary education (Grades 1 to 8) over the same period (MOE 2013). Nevertheless, despite the increases in both the number of schools and the enrolment rate, the performance of children in the Young Lives sample has deteriorated.[1] For instance, in literacy tests administered to the Older Cohort in 2006 and the Younger Cohort in 2013, when each cohort was 12 years old, the proportion of children who were unable to read anything increased from 10 per cent in 2006 to 14 per cent in 2013 (Woldehanna and Pankhurst (2014). These figures signify a serious decline in the quality of education over this period, and students' declining educational performance could prevent the Ethiopian Government from achieving its aims. Therefore there is a need to explore children's level of learning and observe the performance of different socio-economic groups. Furthermore, knowledge about this and evidence of it will be important policy inputs for the second phase of the Growth and Transformation Plan (GTP-II) and the development agenda beyond 2015.

The main objective of this study is to examine children's levels of learning, and the specific aims are (1) to compare the mathematics, Peabody Picture Vocabulary Test (PPVT) and common reading scores of children of the same age in 2006 and 2013 and examine the extent of the changes in levels of learning, and (2) to examine inequality in learning outcomes by assessing the extent of the changes in learning for different socio-economic groups in each of the relevant survey rounds and between rounds.

In light of this, the paper uses statistical analysis to compare the level of learning of 12-year-old children in 2006 (Older Cohort, Round 2) and 2013 (Younger Cohort, Round 4), to see if the levels of learning are declining, using the PPVT, mathematics and the common item reading test results. Then the test scores are disaggregated by gender, urban/rural location, region, caregiver education, and whether children attended private school or government school. Furthermore, the study compares the decline in the levels of learning of these groups and analyses the differences in scores between cohorts and between different characteristics in each cohort by employing standard and popular statistical tests such as student t-test and χ^2 test statistics.

The rest of the paper is organised as follows. Section 2 briefly reviews the general literature and specific literature for Ethiopia on levels of learning and Section 3 presents the nature of the data and the method used in this study. Section 4 discusses the results from the descriptive statistics and, finally, concluding remarks are delivered in Section 5.

1 Young Lives is a unique longitudinal study of childhood poverty tracking the lives of 12,000 children in four countries (Ethiopia, Andhra Pradesh and Telangana in India, Peru and Vietnam). In each country, the survey follows 3,000 children in two age cohorts (a Younger Cohort, born in 2001-02, and an Older Cohort aged 8 in 2001/02). In Ethiopia, the children live in 20 sentinel sites located in five regions of the country (Addis Ababa, Amhara, Oromia, SNNP and Tigray). Young Lives collects information on household and child characteristics, children's well-being (health, nutrition, education), access to services, and education outcomes including scores in a quantitative test, and the Peabody Picture Vocabulary Test (PPVT).

2. Literature review

2.1. General literature on levels of learning

Education plays a key role in the development process of a country, as an educated labour force and skilled workers make economic growth and poverty reduction progress more quickly. According to Psacharopoulos and Patrinos (2004), who worked with data from a wide variety of countries, with each additional year of education, a person's earnings increase on average by about 10 per cent. The study indicated that the returns on schooling were higher for developing countries (11 per cent) than for countries belonging to the OECD (7.5 per cent). Because of global efforts to achieve the second Millennium Development Goal, universal primary education, there has been a substantial increase in the number of children enrolled in school worldwide, in developing countries in particular. However, their learning outcomes vary a great deal between different countries, as well as within the same country over time.

According to the US Department of Education (2000), cited in Denton et al. (2003), in the USA, trends in students' achievements in mathematics and science were characterised by declines in the 1970s, followed by increases during the 1980s and early 1990s, and mostly stable performance after that. Modest gains were also evident in reading, and overall, improvement was most evident in mathematics.

According to Barro and Lee (2010) the average years of schooling in the developing countries increased from 2.0 years in 1950 to 7.2 years in 2010. Beatty and Pritchett (2012), however, argue that progress in learning in many of the developing countries, especially in the areas of literacy, mathematics, and science, was stagnant or even declining. These authors estimated that, at their current rates, Colombia would take 30 years and Turkey would take 194 years to reach the average level of learning for the OECD countries while countries such as Indonesia, Iran, Jordan, Malaysia, Thailand and Tunisia would never catch up, as learning levels had actually declined from one testing period to the next. Besides, they argued that most countries were expected to achieve the Millennium Development Goal of universal primary education but that students accomplished their schooling without mastering basic literacy and numeracy skills.

Pandey et al. (2006) studied learning levels and gaps in Pakistan with a focus on students' achievement as measured through test scores. They found that learning levels were low: few students had mastered mathematics to the appropriate level and only about 30 per cent of the students were able to supply the missing word in a sentence, according to the tests administered at the end of the third grade. Besides, huge educational gaps were also observed between schools; for instance, the gap in English test scores between government and private schools was 12 times the gap between children from rich and poor families. Nevertheless, some issues remained puzzling to the authors of this study. The first puzzle was the variation across government schools; given that there were few differences in performance incentives for government school teachers, it remained mysterious as to why test scores across government schools were so different. The second puzzle was why children in Pakistani villages didn't go to the better government schools in the village, given that there were good and bad government schools in every village and no residential and admission restrictions on which schools children could attend.

Clemens (2004) argued that economic conditions and parental education levels determined children's school enrolment rather than education policy interventions and hence some

countries would take a long time to boost school enrolment. Besides, he argued that there was a remarkable uniformity of experience in the increase in enrolment rates but that the poor countries had raised enrolment by accepting dramatic declines in schooling quality, thereby failing large numbers of students, or by adopting other practices that cast doubt on the sustainability or exportability of their techniques.

It is believed that increasing school enrolment has brought about a deterioration in the quality of education to the extent that most children in low-achieving countries are unable to comprehend texts appropriate to their grade (Aga Khan Foundation 2010).

The Learning Metrics Task Force (2013) in their report *Toward Universal Learning: A Global Framework for Measuring Learning* state that though millions of children have been tracked and brought into the school arena, and awareness regarding education has been created, there remains much work to do regarding how to adequately measure and track success at the global level. And their report provides some insight into how learning should be measured and how to improve the measurement of learning outcomes and thereby improve learning levels themselves.

Murray (2012) studied the factors shaping education inequalities using longitudinal Young Lives data. She argued that educational inequalities among children began before they were enrolled in primary school and never reversed as they got older and this pattern prevailed in all the study countries. The major gaps that the author identified were in relation to geographic location, household poverty, ethnicity/language, levels of parental education and gender. She suggested that countries' development plans should consider the marginalised groups and focus on the specific needs of the different groups, and she asserted that education quality improvement strategies needed to consider the achievement gaps that existed between different groups.

Singh and Sarkar (2012) used Young Lives longitudinal data to investigate whether the 'private school premium' was the result of better teaching in private schools in the state of Andhra Pradesh in India. They found that the mathematics test scores of children in private schools were significantly higher than those of children in government schools and they argued that it was the teacher's activity in the classroom that had an impact on children's learning outcomes, not the experience, gender and other standard characteristics of the teacher or the characteristics of the children themselves. Besides, children taught by teachers with professional qualifications were found to have higher learning outcomes than those taught by teachers with only senior secondary education. Nevertheless, children taught by holders of Bachelors' or Masters' Degrees were not found to have better outcomes. Consequently, they conclude that there is a need for creating appropriate benchmarks for both private and government schools to set standards for teaching and learning.

Furthermore, A. Singh (2013) studied the size and sources of the private school premium in test scores in India and examined the value added by private schools and the sources of learning in these schools using the value-added learning models. The raw test scores for private and government schools were found to be very different. In private schools in rural areas the performances of younger children in English and receptive vocabulary were better and those of older children were found to be better in mathematics and Telugu. In general, there were significant differences in English, heterogeneous differences in Telugu and no significant differences in mathematics for 8- to 10-year-old children; and a significant effect on mathematics, Telugu and receptive vocabulary for 15-year-old children in rural areas. However, no significant effects were found in the urban areas.

Singh (2014), in his study on the emergence and evolution of learning gaps across four countries, using the unique child-level data obtained from Ethiopia, India, Peru, and Vietnam, found that substantial learning gaps exist across these countries with Ethiopia at the lower end and Vietnam at the upper end. And the magnitudes of the learning gap across these countries widens as the children grow up, still with children from Ethiopia at the lower end. The researcher explained this divergence primarily by the differential productivity of a year of schooling and also due to country-specific enrolment guidelines.

Afrassa and Keeves (1999) investigated the levels of mathematics achievement of lower-secondary school Australian students over time using three different datasets and compared three groups of students. The first sample consisted of 13-year-old government school students in Grades 7, 8 and 9 in 1964 from five states; the second sample, from 1978, included students in non-governmental schools and covered seven states; while third sample, from 1994, covered both types of school and all Australian states. The findings revealed that the levels of mathematics achievement of Australian students at the lower-secondary school level had declined over the three decades prior to 1994. Furthermore, they suggested that there was a need to investigate differences in conditions of learning and it was necessary to conduct further research to identify the reasons for the decline.

2.2. Literature on levels of learning in Ethiopia

Over recent decades, the central objectives of the Ethiopian national education strategies have been to address the country's human development needs and achieve the relevant Millennium Development Goals by 2015. Since the vast majority of the people live in rural and dispersed communities, delivering education with equitable access and quality has been a big challenge. The Government of Ethiopia has been undertaking several Education Sector Development Programmes (ESDP I, II, III and IV). The overall goal of the education sector in the ESDP III was to achieve the relevant Millennium Development Goals and to meet the relevant objective of the national development plan through supplying a qualified, trained workforce with the necessary quantity and quality at all levels. Under ESDP III, Ethiopia made significant progress in education. Access, at all levels of the education system, increased at a rapid rate in line with a sharp increase in the number of teachers, schools and institutions. There were significant increases in the availability of trained teachers and some other inputs which are indispensable for a high-quality education system. The primary school net enrolment rate almost doubled in the first decade of the century, rising from just 48 per cent in 2000/1 to 86 per cent in 2012/3 (MOE 2001, 2013). Disparities between more privileged groups and disadvantaged, deprived groups and emerging regions (underdeveloped regions) regarding primary net enrolment rates and the availability of trained teachers declined.

Despite the remarkable progress made towards the Millennium Development Goals and the Education For All goals, disparities in the quality of education and learning outcomes, as well as high grade repetition and drop-out rates, have remained obstacles to the development of the Ethiopian education sector in particular and that of the developing economies in general. It was estimated recently that about 3 million primary-school-age children remained out of school in Ethiopia, among them a significant number from emerging regions (UNICEF 2012).

To improve the learning outcomes of students and thereby successfully pursue its educational strategies, the Government of Ethiopia has been giving due emphasis to concerns about schooling quality. The first phase of the GEQIP, which it launched in 2008, aims to improve the quality of general education throughout the country. Education policies

aimed at improving quality will draw on the GEQIP and further develop the package. The GEQIP will thus become an integral part of ESDP IV (MOE 2012).

Nonetheless, students' achievement declined significantly in the first half of the run-up to 2015; for instance, the composite score for learning achievement in Grade 4 fell from 48 per cent in 2000 to 41 per cent in 2007, though it has remained relatively flat since. The score stood at 40 per cent in 2012, which suggests that the efforts to improve schooling quality may be starting to arrest its decline in the first cycle of primary education. The composite score for Grade 8 declined from 43 per cent to 40 per cent between 2000 and 2007, well below the required levels, and continued to decline since 2007 and reached 35.3 in 2012 (NEAEA 2008a, 2008b, 2003a, 2003b). Another below-par outcome according to the 2010 Ethiopia Early Grade Reading Assessment (EGRA) (RTI International 2010) was in the area of reading. About 80 per cent of students were not able to read with fluency and disparities were also observed by gender and rural/urban location. Further, the national Grade 1 drop-out rate decreased slightly from 28 per cent in 2000/1 to 25 per cent in 2011/2, though it varies across regions.

Tesfay (2012) examined horizontal inequalities in children's educational outcomes in Ethiopia using longitudinal data from Young Lives. Her research explored the effect of slow grade progression on cognitive development between the ages of 12 and 15. From her study one can deduce that there was significant variation in educational achievement between different ethnic groups, with levels being consistently higher in Addis Ababa. The findings show that Hadiyya and Sidama children were significantly behind the official grade for their age in comparison to their peers.

The current research paper investigates the learning levels of children and sees whether it declined between 2006 and 2013. The above literature review suggests that there were declining levels of achievement and deterioration in the quality of education. However, not much study has been done to explore the learning outcomes of students since the launch of the GEQIP in 2008. There is a need to investigate the learning levels of students over time, to see whether they are declining and for whom.

3. Data and method

3.1. Data

The current study uses Younger and Older Cohort data from the survey undertaken by Young Lives in Ethiopia, in which the same children are followed over the course of the study (2002–17). The survey was conducted in four rounds. The first round was carried out in the last quarter of 2002, sampling 1,999 children aged 6–18 months (referred to as the Younger Cohort) and 1,000 children who were 7.5 to 8.5 years old (referred to as the Older Cohort). The Round 2 survey was conducted in 2006, Round 3 in 2009 and the Round 4 in 2013. The regions were selected because 96 per cent of the population of the country lives in these areas. The selection criteria adopted to choose the 20 sentinel sites were that they had to be located in poor areas based on the country's food insecurity designation. Seventy-five per cent of the sentinel sites in each region were selected from high food-deficit *woredas* (districts) while 25 per cent were selected from lower food-deficit *woredas*. Children in rural areas comprised 60 per cent of the sample while 40 per cent were from urban areas. Each region contained 20 per cent of the total sample except for Addis Ababa, which contained 15

per cent of the sample and Southern Nations, Nationalities and People's Region (SNNPR), from which 25 per cent of the sample was selected. Moreover, consultations were made with regional policymakers and other stakeholders to guide the selection of sentinel sites. Within each sentinel site, a simple random sample of 100 households was taken.

The household questionnaires contain questions on livelihood and asset framework, household food and non–food expenditure and economic changes and recent life history. The data used for the analysis are mostly obtained from this part of the questionnaire in addition to the household demography variables.

3.2. Method

The Older Cohort children were 12 years old in Round 2 (2006) and the Younger Cohort children were 12 years old in Round 4 (2013). The study uses data from Round 2 for the Older Cohort and Round 4 for the Younger Cohort to compare the PPVT, maths test, and reading test scores of the sampled children. These two sets of data are used to compare the levels of learning among children of the same age seven years apart.

The study uses simple descriptive statistics to compare these levels of learning. Learning outcomes were measured in three different ways. First, children's literacy was measured by a reading test. Children were shown literacy cards and asked to read the letters, words and sentences on it. The test was conducted in their mother tongue. We disaggregated the test results by gender, region, type of school (private or government), location (urban or rural) and household wealth level. As the reading test outcomes are categorical, a chi-square test was used to see if there was a statistically significant relationship between these outcomes and different child characteristics both within the groups and between them.

Secondly, for the maths test, the children were given a test paper and left to work by themselves unless they asked a question on how to answer. Regarding the similarity of the test between rounds, the number of questions in the maths test increased in Round 4. However, in spite of this increase, the content remained similar to that in Round 2. In the same fashion the learning outcome obtained from the maths test is assessed across different groups. Then an independent t-test (also referred to as an independent-samples t-test, independent-measures t-test or unpaired t-test) is used to determine whether the mean of a maths percentage raw score was the same in two unrelated, independent groups; for instance, male versus female or urban versus rural. However, when we have three or more independent and unrelated groups, the one-way analysis of variance (ANOVA) is used. For example, a one-way ANOVA can be used to determine whether the mean maths scores differed among children from different regions. However, to determine which specific groups are significantly different from the others, there is a need to conduct post hoc tests. There are many types of post hoc test that one can use following a one-way ANOVA, for instance, Bonferroni, Sidak, Scheffe, Tukey, etc.; in this paper the Tukey post hoc test was used.

Finally we compare children's PPVT scores. The PPVT measures the overall educational attainment of children. The test requires respondents to select the pictures that best represent the meaning of a series of stimulus words read out by the examiner. It consists of 17 sets of 12 words each and the raw score test results can take possible values from 0 to 204. As with the maths test, an independent t-test is used to determine whether the mean of PPVT percentage raw scores is the same in two unrelated, independent groups. However, when we have three or more independent and unrelated groups, the one-way ANOVA is used and to determine which specific groups are significantly different from the others, post hoc tests are also conducted.

4. Results and discussion

In this section the results from the descriptive statistics are presented and discussed. In the first sub-section the performance in reading of the two groups of 12-year-olds is assessed and then disaggregated by different socio-economic characteristics in order to see who is affected more. The next two sub-sections present the maths and PPVT scores, and Section 4.4. discusses all the results.

4.1. Reading outcomes

The results of the reading test fell into four categories: cannot read anything, can read letters, can read words, and can read sentences. We then compared the test outcomes of the two cohorts and investigated whether differences in the type of school, area of residence, and gender were correlated with significant differences in these outcomes, before investigating the scores of children in the same cohort, again disaggregating the results in the same way.

4.1.1. Inter-cohort comparison of reading outcomes

Table 1 presents the percentage of 12-year-olds who could not read anything in both rounds. We can see that the percentage of such children increased from 10 per cent in 2006 to 14 per cent in 2013. Furthermore, even though the percentage of children who could read sentences was higher in 2013 than in 2006 (see Tables 2 and 3), the total level of reading (including letters and words) had declined from 90 per cent in 2006 to 86 per cent in 2013 (Tables 2 and 3). This shows that the level of reading of the sampled 12-year-olds showed a deteriorating trend between 2006 and 2013. Moreover, the t-test shows that the mean difference between the percentages of those who could not read anything in the two rounds was statistically significant at 1 per cent level of significance, with a mean difference of 3.87.

Table 1 also shows that the percentage of boys who could not read anything increased from 10.75 per cent in 2006 to 13.75 per cent in 2013; the proportion of girls who could not read anything also showed an increasing trend. However, the rate of decline in reading is higher for girls than boys but the t-test results indicate that the difference between the average scores of boys between the rounds was significant at the 5 per cent level of significance and between girls was also significant at the 10 per cent level of significance. Table 1 also gives us a picture of how the reading outcomes of children changed between the two periods disaggregated by the caregiver's education level. The percentage of children unable to read increased for all groups except those whose caregiver had attended adult literacy or religious education. The t-test results indicate that the difference in the percentage of children who could not read and whose caregiver had had no education was found to be statistically significant at 1 per cent level of significance. However, the difference in the percentage of children who could not read anything between the two rounds was found to be statistically insignificant for all the levels of caregiver's education except 'none'.

Table 1. *Percentage of 12-year-olds who could not read anything in 2006 and 2013 and mean differences*

	Older Cohort, age 12 (2006 – R2)	Younger Cohort, age 12 (2013 – R4)	Difference (R4–R2)	Percentage difference [(R4–R2)/R2*100]
Gender				
Female	9.39	14.14	4.82**	50.59
Male	10.57	13.66	3.09*	29.23
Caregiver's education				
None	11.70	18.84	7.14***	61.03
Adult literacy or religious	7.84	7.32	-0.53	-6.63
Lower primary (1–4)	9.74	12.50	2.76	28.34
Upper primary (5–8)	10.94	11.67	0.73	6.67
Grade 8 and above	2.67	4.50	1.84	68.54
Urban or rural site				
Rural	12.67	18.06	5.39***	42.54
Urban	6.06	7.89	1.83	30.20
Region				
Tigray	9.45	6.54	-2.91	-30.79
Amhara	8.85	8.45	-0.41	-4.52
Oromia	10.00	13.52	3.52	35.20
SNNPR	15.98	31.53	15.55***	97.31
Addis Ababa	2.10	1.87	-0.23	-10.95
Type of school				
Private	5.26	3.54	-1.72	-32.70
Community	5.88	14.29	8.40	143.03
Government	8.39	11.93	3.53***	42.19
Wealth tercile				
Bottom	14.94	22.38	7.44***	49.80
Middle	9.85	12.68	2.83	28.73
Top	5.21	6.07	0.85	16.51
Total	10.00	13.87	3.87***	38.70
Sample size	980	1875		

Note: *** p<0.01, ** p<0.05, * p<0.1
Source: own computation.

We can observe that both the urban and the rural children have shown declining levels of reading. The reading test outcome of rural children was lower than for urban children in 2006 and it becomes worse in 2013. However, the t-test results indicate that the difference in the percentage of children who could not read anything between the two rounds was found to be statistically insignificant for urban children while it was statistically significant at 1 per cent level for rural children.

The reading test outcome is also disaggregated by the geographic regions the children live in. As can be seen in Table 1, levels of reading have increased for some of the regions (Tigray, Amhara and Addis Ababa) and decreased for others (Oromia and SNNPR). The steepest decline in reading outcomes is observed in SNNPR, where the percentage of children who could not read anything almost doubled, and this difference was also statistically significant at the 1 per cent level of significance. An increase in the proportion of

children who could not read anything is also observed in Oromia but it was found to be statistically insignificant.

Table 1 also presents the percentage of non-readers disaggregated by the type of school they were attending. It shows that the percentage of children who could not read anything decreased among children attending private schools but it increased for children attending government and community-funded schools (supported by NGOs, charity or religious organisations). For government-funded school attendees the difference was found to be statistically significant at the 1 per cent level of significance.

The learning differences were also disaggregated by wealth terciles.[2] The wealth terciles were calculated from the Round 2 wealth index for the Older Cohort children and from the Round 3 wealth index for the Younger Cohort children. It was found that the percentage of children who could not read anything had increased for all wealth terciles but this increase was only statistically significant for the bottom tercile, and the percentage difference was also the highest for this group (i.e. about 50 per cent more children were unable to read in 2013 than in 2006.

Table 1 also presents the percentage difference between the rounds and indicates that the highest percentage difference was observed for community school attendees (with non-readers increasing by 143 per cent) followed by children from SNNPR (with an increase of 97 per cent). On the other hand, the percentage of children who could not read decreased for private school attendees by 32 per cent, for children from Tigray by 31 per cent, from Addis Ababa region by 11 per cent, and from Amhara by 5 per cent, and for children whose caregiver had received adult literacy or religious education by about 7 per cent.

4.1.2. Intra-cohort comparison of reading outcomes

We also analysed the reading outcomes in each round by different child and household characteristics. Since the reading outcomes are categorical, we used a chi-square test to assess if there was a relationship between the reading outcomes and different child and household characteristics.

Table 2 presents the results of the chi-square tests of the outcomes of the children who were 12 years old in 2006. The chi-square test reveals that there was no significant difference between boys and girls in reading outcomes, meaning that there was no statistically significant relationship between the reading outcomes and gender (chi-square with three degree of freedom = 4.2483, p = 0.236). However, the chi-square test results suggest that there was a statistically significant relationship between reading outcomes and caregiver's education (chi-square with three degree of freedom = 27.9798, p = 0.006). One can observe that the proportion of children who could not read anything was highest for children whose caregivers had had no education, followed by those whose caregivers had been educated to lower primary level, and those whose caregivers had received adult literacy or religious education. Similarly, a statistically significant relationship was found between the reading outcomes of children and their residence in urban or rural sites. Children who lived in rural sites performed worse than children in urban areas, and the difference was found to be statistically highly significant (with Pearson chi2(3) = 40.8810, P = 0.000). There was also a statistically significant difference between reading outcomes and the regions in which

2 The wealth index is constructed from housing quality, consumer durables and access to services, including electricity, clean water, sanitation and cooking fuel.

children lived, at the 1 per cent level of significance. And, a statistically significant difference among the types of schools was also found, though at the 10 per cent level of significance. Moreover, it was also observed that there were statistically significant differences among the wealth tercile groups.

Table 2. *Reading outcomes of children in 2006 (Chi-square results)*

	Can't read anything	Reads letters	Reads words	Reads sentences	Test statistics (d.f) (P-vale)
Gender					
Male	10.75	15.62	16.02	57.61	Pearson chi2(3) =4.2483
Female	9.47	13.47	13.05	64.00	P-value = 0.236
Difference	1.28	2.15	2.97	-6.39	
Caregiver's education					
None	11.93	17.14	15.62	55.31	Pearson chi2(12) = 27.9798
Adult literacy or religious	7.95	11.92	14.57	65.56	P-value = 0.006
Lower primary (1–4)	10.32	13.17	14.59	61.92	
Upper primary (5–8)	2.70	9.46	6.76	81.08	
Grade 8 and above	0.00	0.00	100.00	0.00	
Urban or rural site					
Urban	6.09	9.64	11.68	72.59	Pearson chi2(3) = 40.8810
Rural	12.89	17.94	16.55	52.61	P-value = 0.000
Difference	-6.8	-8.3	-4.87	19.98	
Region					
Tigray	9.45	19.40	13.93	57.21	Pearson chi2(12) = 49.3697
Amhara	9.14	8.06	11.83	70.97	P-value = 0.000
Oromia	10.15	13.71	20.30	55.84	
SNNPR	16.18	19.09	11.20	53.53	
Addis Ababa	2.10	9.79	16.78	71.33	
Type of school					
Private	5.26	2.63	10.53	81.58	Pearson chi2(9) = 15.0848
Community	5.88	5.88	5.88	82.35	P-value = 0.089
Government	8.44	15.28	15.16	61.12	
Wealth tercile					
Bottom	15.22	18.63	16.15	50.00	Pearson chi2(6) = 36.2889
Middle	10.03	15.05	13.48	61.44	P-value = 0.000
Top	5.21	10.12	14.11	70.55	
Total	10.12	14.57	14.57	60.74	

Note: *** p<0.01, ** p<0.05, * p<0.1
Source: own computation.

The chi-square tests were also conducted on the results of the children who were 12 years old in 2013 (see Table 3). As regards differences between social groups, Table 3 portrays results similar to those obtained in 2006. The difference in reading outcomes was statistically insignificant for both boys and girls in the sample. However, one can see that there was a statistically significant relationship between reading outcomes and other characteristics: caregiver's education, urban or rural site, region, type of school and wealth tercile. These were all statistically significant at the 1 per cent level.

Table 3. *Reading outcomes of children in 2013 (Chi-square results)*

	Can't read anything	Reads letters	Reads words	Reads sentences	Test statistics (d.f) (P-vale)
Gender					
Male	13.75	10.49	10.79	64.97	Pearson chi2(3) = 1.0012
Female	14.29	9.14	11.09	65.49	P-value = 0.801
Difference	-0.54	1.35	-0.30	-0.52	
Caregiver's education					
None	19.15	11.11	10.99	58.75	Pearson chi2(12) = 68.7449 P-value = 0.000
Adult literacy or religious	7.41	6.79	9.26	76.54	
Lower primary (1–4)	12.12	10.85	11.96	65.07	
Upper primary (5–8)	5.99	4.79	9.58	79.64	
Grade 8 and above	0.00	3.64	7.27	89.09	
Urban or rural site					
Urban	7.93	4.94	5.98	81.14	Pearson chi2(3) = 147.1786
Rural	18.29	13.33	14.43	53.95	P-value = 0.000
Difference	-10.36	-8.39	-8.45	27.19	
Region					
Tigray	6.58	7.89	4.21	81.32	Pearson chi2(18) = 408.5247 P-value = 0.000
Amhara	8.54	4.13	7.99	79.34	
Oromia	13.55	16.88	20.20	49.36	
SNNPR	32.30	14.16	15.27	38.27	
Addis Ababa	1.87	3.00	3.37	91.76	
Type of school					
Private	3.54	2.65	5.31	88.50	Pearson chi2(12) = 26.3587 P-value = 0.010
Community	14.29	14.29	11.43	60.00	
Government	12.01	10.02	11.39	66.58	
Wealth tercile					
Bottom tercile	22.82	14.72	13.92	48.54	Pearson chi2(6) = 146.1498
Middle tercile	12.78	9.39	10.36	67.48	P-value = 0.000
Top tercile	6.08	5.42	8.37	80.13	
Total	**14.00**	**9.85**	**10.93**	**65.21**	

Note: *** p<0.01, ** p<0.05, * p<0.1
Source: own computation.

4.2. Mathematics test outcomes

The results of the mathematics tests were analysed a similar way to those of the reading test.

4.2.1. Inter-cohort comparison of maths test scores

Table 4 presents the mean percentage of correct answers in the maths tests in Round 2 for the Older Cohort and Round 4 for the Younger Cohort. It shows that the average score of children in 2006 was 54.4 per cent while in 2013 it was 37.2 per cent. The gap between the average maths test scores in 2006 and 2013 is very wide; there was a percentage gap of about 17.3 per cent between the two rounds and the t-test results also revealed that the difference was statistically significant at 1 per cent level of significance.

Table 4. *Mean maths raw scores in 2006 and 2013 (%)*

	Older Cohort, age 12 (2006 – R2)	Younger Cohort, age 12 (2013 – R4)	Difference (R4-R2)	Percentage difference [(R4-R2)/R2*100]
Gender				
Female	53.05	37.32	-15.74***	-29.65
Male	55.76	37.09	-18.67***	-33.48
Caregiver's education				
None	49.69	31.98	-17.71***	-35.64
Adult literacy or religious	52.18	34.74	-17.44***	-33.42
Lower primary (1–4)	57.91	34.55	-23.36***	-40.34
Upper primary (5–8)	62.12	43.26	-18.85***	-30.36
Grade 8 and above	67.11	51.39	-15.72***	-23.42
Urban or rural site				
Rural	48.04	28.05	-19.99***	-41.61
Urban	63.42	48.58	-14.85***	-23.40
Region				
Tigray	61.47	33.15	-28.32***	-46.07
Amhara	49.72	34.73	-14.99***	-30.15
Oromia	45.80	33.97	-11.83***	-25.83
SNNPR	49.74	32.17	-17.57***	-35.32
Addis Ababa	69.93	56.57	-13.36***	-19.10
Type of school				
Private	73.68	61.74	-11.95***	-16.21
Community	73.20	46.34	-26.86***	-36.69
Government	53.57	35.96	-17.61***	-32.87
Wealth tercile				
Bottom	46.24	26.94	-19.30***	-41.73
Middle	51.40	34.04	-17.36***	-33.77
Top	65.03	48.55	-16.48***	-25.35
Total	54.42	37.17	-17.25***	-31.70
Sample size	950	1638		

Note: *** p<0.01, ** p<0.05, * p<0.1
Source: own computation.

Disaggregating the results by gender shows that the rate of decline is higher for boys (18.7 percentage points) than girls in the sample (15.7 percentage points). All regions have seen a declining trend in the average raw maths scores. The highest rate of decline is observed in Tigray (28.3 percentage points), followed by SNNPR (17.6), while the lowest is observed in Oromia (11.8 percentage points), followed by Addis Ababa (13.4).

The table also shows that the test results of children in rural areas have higher rates of decline than those in urban areas. The average score declined by 20 percentage points for rural children (from 48 per cent in 2006 to 28 per cent in 2013) while that of urban children declined by 15 percentage points (from 63.4 per cent in 2006 to 48.6 per cent in 2013). The mean difference between the rounds was also found to be statistically significant for both rural and urban children.

Children attending private school see a lower rate of decline though their average scores decreased from 73.68 per cent to 61.74 per cent, whereas children attending community schools have the highest rate of decline, with a fall of about 27 percentage points. Children attending government-funded schools have seen their test scores fall by about 18 percentage points. The mean difference between the rounds was found to be statistically significant for all categories. When the results are analysed by caregiver's education level, children from all groups have seen declining scores and the differences were also statistically significant. Disaggregation by wealth tercile reveals that the scores of children from in the bottom tercile showed the steepest rate of decline (about 42 per cent), also statistically significant at the 1 per cent level of significance, whereas those in the top wealth tercile have declined the least (25 per cent).

In general, from the results, one can see that maths test scores have declined for all groups. The percentage differences were the highest for children from Tigray (the maths score has declined by about 46 per cent) followed by children from rural areas (42 per cent) and children whose caregivers had been educated to lower primary level (40 per cent); the lowest percentage differences, however, were observed for children attending private school (16 per cent) and from Addis Ababa (19 per cent). Hence, there is deterioration in the levels of learning in maths between 2006 and 2013 and it is also statistically significant.

4.2.2. Intra-cohort comparisons of maths test scores

In this part, t-tests and the one-way ANOVA were used to determine whether the mean of the maths raw score was statistically the same for two or more unrelated, independent groups. Table 5 presents the average maths raw scores in each round for different independent groups. Accordingly, the difference in mean maths raw score was found to be statistically insignificant for boys and girls in both 2006 and 2013 though the difference was positive in 2006 and negative in 2013, meaning that boys outperformed girls in 2006 but girls outperformed boys in 2013. However, whether children lived in an urban or a rural site brought about a significant difference in the average maths raw scores of the 12-year-olds in both 2006 and 2013. In 2006 the mean difference between urban and rural was about 15.38 percentage points but the figure rose to 20.52 percentage points in 2013; this indicates that there is a huge and significant gap between children from rural and urban sites. Hence, we can deduce that rural children are lagging behind as they show a higher declining rate and their recent performance is far less good than that of their urban counterparts.

Table 5. *Mean difference of maths raw scores in 2006 and 2013 (horizontal comparison)*

	Older Cohort, age 12 (2006 – R2)	Younger Cohort, age 12 (2013 – R4)
Gender		
(Male) – (Female)	2.71	-0.23
Urban or rural site		
(Urban) – (Rural)	15.38***	20.52***
Caregiver's education		
(Adult literacy or religious) - (None)	2.49	2.75
(Lower primary (1–4) – (None)	8.22***	2.57
(Upper primary (5–8)) – (None)	12.43***	11.28***
(Grade 8 and above) – (None)	17.42***	19.41***
(Lower primary (1–4)) – (Adult literacy or religious)	5.73	-0.18
(Upper primary (5–8)) – (Adult literacy or religious)	9.94**	8.53***
(Grade 8 and above) – (Adult literacy or religious)	14.93***	16.66***
(Upper primary (5–8)) – (Lower primary (1–4))	4.21	8.71***
(Grade 8 and above) – (Lower primary (1–4))	9.20	16.84***
(Grade 8 and above) – (Upper primary (5–8))	4.99	8.13***
Region		
(Amhara) – (Tigray)	-11.75***	1.58
(Oromiya) – (Tigray)	-15.67***	0.82
(SNNPR) – (Tigray)	-11.74***	-0.99
(Addis Ababa) – (Tigray)	8.46**	23.42***
(Oromiya) – (Amhara)	-3.92	-0.76
(SNNPR) – (Amhara)	0.02	-2.57
(Addis Ababa) – (Amhara)	20.21***	21.84***
(SNNPR) – (Oromiya)	3.93	-1.81
(Addis Ababa) – (Oromiya)	24.13***	22.60***
(Addis Ababa) – (SNNPR)	20.19***	24.41***
Type of school		
(Community) – (Private)	-0.48	-15.39***
(Government) – (Private)	-20.11***	-25.78***
(Government) – (Community)	-19.63**	-10.38**
Wealth tercile		
(Middle) – (Bottom)	5.17**	7.10***
(Top) – (Bottom)	18.79***	21.61***
(Top) – (Middle)	13.63***	14.51***

Note: *** p<0.01, ** p<0.05, * p<0.1
Source: own computation.

Regarding the relationship between mean maths score and caregiver's education, one can see that the differences in the caregiver's education are correlated with significant differences in the average maths raw scores. Similarly, a significant difference among regions was also observed. Compared to Addis Ababa, all the regions performed significantly poorly, with SNNPR faring worst. Furthermore, the type of school the children attended was also found to be correlated with significant differences in their scores. The difference between children attending government-funded and private schools was about 20.11 percentage points in

2006 and rose to 25.78 percentage points in 2013 and was also statistically significant at 1 per cent level of significance.

The difference in mean maths raw score was found to be statistically significant for all the wealth tercile groups in both 2006 and 2013, and in 2013 the percentage differences between the groups had increased and all were statistically significant at 1 per cent level of significance. Hence, we can deduce that there is a statistically significant difference between those in the top and bottom wealth terciles, with children at the top performing better than those in the bottom and middle wealth terciles.

Furthermore, we also calculated the percentage of correctly answered questions for each type of question in both rounds in the Appendix (see Table A4) and found that the average percentages in all the question types were lower for the Round 4 children than for their Round 2 counterparts. This implies that there is no room left for the difference in test results to be due to test differences; i.e. the difference in test scores between the cohorts may not be attributed to the difference in tests.

4.3. PPVT score outcomes

4.3.1. Inter-cohort comparisons of PPVT scores

The PPVT is a widely used test of receptive vocabulary intended to provide a quick estimate of verbal ability and scholastic aptitude. It was developed in 1959 by special education researchers Lloyd M. Dunn and Leota M. Dunn. The test requires respondents to select the pictures that best represent the meaning of a series of stimulus words read out by the examiner.

Table 6 shows the average percentage of children with correct PPVT raw scores in 2006 and 2013. We can see that, overall, the mean percentage of corrected raw scores in the PPVT increased by 2.41 percentage points and was statistically significant at the 5 per cent level of significance, showing that the performance of the 12-year-olds in the PPVT improved between 2006 and 2013. When the results are disaggregated by gender we can see that the scores of both boys and girls increased. The difference between the scores in the two survey rounds was found to be statistically significant for boys but not for girls. Disaggregating the scores by urban/rural location reveals that urban children's scores have increased by about 7 percentage points and that the mean difference was statistically significant. However, the scores of children from rural areas have declined, although the difference between the rounds was statistically insignificant. Moreover, not only have children from rural areas seen their scores decline but also the gap between their average score and that of urban children has become much wider. In 2006, this gap was about 17 percentage points, but it increased to about 25 percentage points in 2013.

Table 6. *Average percentage of corrected PPVT raw scores in 2006 and 2013 (vertical comparison)*

	Older Cohort, age 12 (2006 – R2)	Younger Cohort, age 12 (2013 – R4)	Difference (R4-R2)	Percentage difference [(R4-R2)/R2*100]
Gender				
Female	61.31	63.49	2.17	3.56
Male	62.14	64.87	2.72*	4.39
Caregiver's education				
None	56.81	61.66	4.85***	8.54
Adult literacy or religious	63.21	65.70	2.49	3.94
Lower primary (1-4)	64.43	60.85	-3.58	-5.56
Upper primary (5-8)	67.37	63.11	-4.25	-6.32
Grade 8 and above	74.32	78.70	4.38	5.89
Urban or rural site				
Rural	54.49	53.74	-0.75	-1.38
Urban	72.41	78.98	6.57***	9.07
Regions				
Tigray	62.28	69.62	7.34***	11.79
Amhara	54.64	64.5	9.86***	18.05
Oromia	59.80	75.43	15.63***	26.14
SNNPR	57.90	36.04	-21.86***	-37.75
Addis Ababa	79.76	87.82	8.05***	10.11
Type of school				
Private	80.37	87.47	7.10***	8.83
Community	75.03	69.19	-5.84	-7.78
Government	60.53	63.43	2.90***	4.79
Wealth tercile				
Bottom	53.63	49.54	-4.09**	-7.62
Middle	58.84	62.50	3.66**	6.22
Top	72.62	81.05	8.43***	11.61
Total	**61.74**	64.15	2.41**	3.90
Sample size	**964**	**1,872**		

Note: *** p<0.01, ** p<0.05, * p<0.1
Source: own computation.

If we look at children's mean scores disaggregated by their caregivers' education levels, we can see that the scores increase with the caregiver's level of education in Round 2 but not in Round 4. However, comparing the scores for the two rounds, the increasing rate is the highest and statistically significant for children whose caregivers had had no education.

Analysing the results by the type of school the children attended, we can see that the gap between children attending private schools and those in government-funded schools increased. In 2006, this gap was about 20 percentage points but in 2013 it had increased to 24 percentage points. However, the average PPVT score increased slightly between the two rounds and the difference was statistically significant for both private and government-funded school attendees. When analysed by region, the results reveal that the average corrected PPVT raw scores of children living in all regions except SNNPR increased and that the increases were statistically significant. The scores of children from the SNNPR have deteriorated, being about 22 percentage points lower than those of the previous cohort of children at the same age, and the difference between the rounds was found to be statistically significant at the 1 per cent level of significance. The results for the wealth terciles reveal that

the average corrected PPVT raw scores decreased by about 7.6 per cent for children in the bottom wealth tercile, which was statistically significant at 5 per cent level of significance; whereas it increased for children in both the middle and top wealth terciles, indicating a growing inequality of learning outcomes between children from poor and rich families.

4.3.2. Intra-cohort comparisons of PPVT scores

Table 7 presents the relationships between the average PPVT raw scores and different child and household characteristics. It reveals similar results to the analysis of the maths scores (in Table 5), in that there was found to be an insignificant relationship between the gender of children and their mean raw score. However, there is a huge difference between the average PPVT scores of children living in urban sites and those in rural ones. Except in a few of the categories, caregiver's education level was found to be correlated with significant differences in average PPVT scores. There was also found to be a statistically significant relationship between the regions and mean scores in PPVT. The type of school children attended was also found to have a statistically significant relationship with the mean PPVT raw scores of children in both rounds, with the gap between children attending government-funded schools and those going to private schools higher in Round 4 (2013). Moreover, there were big differences between the average PPVT scores of children in different wealth terciles, and the differences among the terciles were found to be statistically significant in both rounds.

Table 7. *Mean difference in PPVT raw scores in 2006 and 2013 (horizontal comparison)*

	Older Cohort, age 12 (2006 – R2)	Younger Cohort, age 12 (2013 – R4)
Gender		
(Male) – (Female)	0.83	1.38
Urban or rural site		
(Urban) – (Rural)	17.92***	25.24***
Caregiver's education		
(Adult literacy or religious) – (None)	6.40***	4.04
(Lower primary (1–4) – (None)	7.62***	-0.81
(Upper primary (5–8)) – (None)	10.56***	1.45
(Grade 8 and above) – (None)	17.51***	17.04***
(Lower primary (1–4)) – (Adult literacy or religious)	1.22	-4.85
(Upper primary (5–8)) – (Adult literacy or religious)	4.15	-2.58
(Grade 8 and above) – (Adult literacy or religious)	11.11***	13.00***
(Upper primary (5–8)) – (Lower primary (1–4))	2.93	2.26
(Grade 8 and above) – (Lower primary (1–4))	9.89***	17.85***
(Grade 8 and above) – (Upper primary (5–8))	6.95*	15.58***
Region		
(Amhara-) – (Tigray)	-7.64***	-5.12**
(Oromiya-) – (Tigray)	-2.48	5.81***
(SNNP) – (Tigray)	-4.38*	-33.58***
(Addis Ababa) - (Tigray)	17.48***	18.20***
(Oromiya) – (Amhara)	5.16**	10.93***
(SNNP) – (Amhara)	3.26	-28.46***
(Addis Ababa) – (Amhara)	25.12***	23.32***
(SNNP) – (Oromiya-)	-1.90	-39.39***
(Addis Ababa) – (Oromiya)	19.97***	12.39***
(Addis Ababa) – (SNNP)	21.86***	51.78***

	Older Cohort, age 12 (2006 – R2)	Younger Cohort, age 12 (2013 – R4)
Type of school		
(Community) – (Private)	-5.34	-18.27***
(Government) – (Private)	-19.84***	-24.04***
(Government) – (Community)	-14.50***	-5.77
Wealth tercile		
(Middle) – (Bottom)	5.21***	12.95***
(Top) – (Bottom)	18.99***	31.51***
(Top) – (Middle)	13.78***	18.55***

Note: *** p<0.01, ** p<0.05, * p<0.1
Source: own computation.

4.4. Discussion of the main findings

The descriptive analysis revealed that the learning level of children was found to be more affected by differences in the household and institutional characteristics than by child characteristics such as gender. The difference in learning outcomes was found to be insignificant for gender. This might be due to the balance in the enrolment and participation of girls in education. For instance, Table 2 indicates that the primary-level Gender Parity Index, which is the ratio of female to male gross enrolment, is close to 1, implying that boys and girls have an equal chance of accessing school. Among the household characteristics, the education level of the caregiver and household's wealth level were found to be correlated with the learning outcomes of children. Location of residence (rural or urban and region) also appeared to make a difference to the learning outcomes of children. Furthermore, from the evidence presented above, the type of school that the children attended was also correlated with the levels of learning of children in both cohorts. Those attending private schools performed better than those attending government schools, maybe because private schools provide better-quality education. A study by Seboka (2003) argued that private schools in Ethiopia emerged primarily in response to the high demand for high-quality education by parents, and hence parents preferred private schools to government schools because of their quality and the low level of bureaucracy involved.

The study has also revealed that educational outcomes, in the areas of literacy and mathematics, deteriorated over the seven years separating the two cohorts of children, and the decline in levels of learning was found to be higher for children from rural areas, those whose caregiver had had little no education, those attending government schools and those from households with lower wealth levels. These phenomena may be explained by the increased enrolment and intake rates in the country. Between 2006 and 2013, the apparent intake rate[3] increased from 124 per cent to 144 per cent and the net intake rate increased from 63 per cent to 96 per cent; the gross enrolment rate increased from 92 per cent to 95 per cent and the net enrolment rate increased from 79 per cent to 86 per cent (see Table A2). However, the school drop-out and grade repetition rates increased from 12.4 per cent and 6.1 per cent to 16.1 per cent and 7.9 per cent respectively (Table A3). Therefore, the deterioration might partially be explained by the high enrolment and intake rates without the necessary equipment and qualified personnel available, which is reflected by the increasing grade repetition and drop-out rates, and ultimately the educational outcomes of children.

3 Apparent intake rate refers to the percentage of new entrants (irrespective of age) in Grade 1 out of the total number of children of the official primary admission age (age 7) in a given year.

5. Conclusion

The paper used Younger and Older Cohort data from the survey undertaken by Young Lives in Ethiopia and employed simple descriptive statistics to compare the levels of learning of two cohorts of children at the same age (12 years old), and also applied t-test, chi-square test and one-way ANOVA tests. Children's reading outcomes have fallen and girls (as compared to boys), children from SNNPR and Oromia as compared to those from other regions in the sample, rural children, children attending community schools, and children with uneducated caregivers have experienced a higher rate of decline in their reading levels. Analysis of children's mathematics test scores also confirms that there is substantial and statistically significant decline in mathematics competency. However, comparing children in the two rounds overall, the average PPVT raw score has increased. But rural children and children from SNNPR have shown a declining trend even in PPVT raw scores and inequality between the private school and government school attendees, and rural and urban children has also been observed; inequality of learning outcome between children from different socio-economic and regional groups, already apparent in 2006, had increased markedly by 2013. Moreover, the chi-square and t-test results also reveal that urban or rural location of residence, caregiver's education, region, type of school, and wealth tercile were found to have significant relationships with the reading and mathematics outcomes, and receptive vocabulary test (PPVT) scores, while gender differences were found to be statistically insignificant. Hence, there is a need to halt the increasing learning inequality among children of the same age but different socio-economic characteristics. As the different socio-economic groups may have differential access to in-school and out-of-school resources that are inputs for education, there should be a public education programme that can reduce the imbalances among children in accessing the inputs required; as appropriate measures could include (1) education grants for children from disadvantaged groups, and (2) equipping government schools and rural schools with the necessary infrastructure and staff to maintain the level of service quality required. A more comprehensive study is crucial in order to identify the in-school and out-of-school factors that affect the quality of learning in mathematics and reading, in order to inform the development of research-based intervention programmes.

References

Afrassa, T. and J. Keeves (1999) 'Changes in Students' Mathematics Achievement in Australian Lower Secondary Schools Over Time', *International Education Journal* 1.1: 1–21.

Aga Khan Foundation (2010) *Improving Learning Achievement in Early Primary in Low-Income Countries: A Review of the Research*, Geneva: Aga Khan Foundation.

Barro, R.J. and J.W. Lee (2010) *A New Data Set of Educational Attainment in the World, 1950–2010*, NBER Working Paper 15902, Cambridge, MA: National Bureau of Economic Research.

Beatty, A. and L. Pritchett (2012) *From Schooling Goals to Learning Goals: How Fast can Student Learning Improve?*, CDG Policy Paper 012, Washington, DC: Center for Global Development.

Clemens, M.A. (2004) *The Long Walk to School: International Education Goals in Historical Perspective*, Working Paper 37, Washington, DC: Center for Global Development.

Denton, K., J. West and J. Walston (2003) *Reading: Young Children's Achievement and Classroom Experiences*, NCES 2003–070, Washington, DC: US Department of Education, National Center for Education Statistics.

LMTF (Learning Metrics Task Force) (2013) *Toward Universal Learning: A Global Framework for Measuring Learning*, Report No. 2 of the Learning Metrics Task Force, Montreal and Washington: UNESCO Institute for Statistics and Center for Universal Education at the Brookings Institution.

MOE (2013) *Education Statistics Annual Abstract: 2005 EC (2012/2013)*, Addis Ababa: Ministry of Education.

MOE (2012) *General Education Quality Improvement Project Phase 2: Project Document*, Addis Ababa: Ministry of Education.

MOE (2010) *Education Sector Development Program IV (ESDP IV)*, Addis Ababa: Ministry of Education.

MOE (2006/07) *Education Statistics Annual Abstract: 1999 EC /2006-07*, Addis Ababa: Ministry of Education.

MOE (2001) *Education Statistics Annual Abstract: 1993 EC /2000-01/*, Addis Ababa: Ministry of Education.

Mullis, I.V.S., M.O. Martin, E.J. Gonzalez and S.J. Chrostowski (2004) *Findings from IEA's Trends in International Mathematics and Science Study at the Fourth and Eighth Grades*, Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Murray, H. (2012) *Is School Education Breaking the Cycle of Poverty for Children? Factors Shaping Education Inequalities in Ethiopia, India, Peru and Vietnam*, Policy Paper 6, Oxford: Young Lives.

NEAEA (2008a) *Ethiopian Second Learning Assessment for Grade 4* (2007), Addis Ababa: National Education Assessment and Examinations Agency.

NEAEA (2008b) *Ethiopian Second Learning Assessment for Grade 8* (2007), Addis Ababa: National Education Assessment and Examinations Agency.

NEAEA (2003a) *Ethiopian Second Learning Assessment for Grade 4* (2002), Addis Ababa: National Education Assessment and Examinations Agency.

NEAEA (2003b) *Ethiopian Second Learning Assessment for Grade 8* (2002), Addis Ababa: National Education Assessment and Examinations Agency.

Pandey, P., T. Zajonc and J. Das (2006) *Learning Levels and Gaps in Pakistan*, Policy Research Working Paper 4067, Washington, DC: World Bank.

Psacharopoulos, G. and H. Patrinos (2004) 'Returns to Investment in Education: A Further Update', *Education Economics* 12.2: 111–34.

RTI International (2010) *Ethiopia Early Grade Reading Assessment. Data Analysis Report: Language and Early Learning*, Addis Ababa: USAID, Ethiopia.

Seboka, B. (2003) *School Choice and Policy Response: A Comparative Context between Private and Public Schools In Urban Ethiopia*, Addis Ababa, Ethiopia.

Singh, A. (2014) *Emergence and Evolution of Learning Gaps across Countries: Panel Evidence from Ethiopia, India, Peru and Vietnam*, Working Paper 124, Oxford: Young Lives.

Singh, A. (2013) *Size and Sources of the Private School Premium in Test Scores in India*, Working Paper 98, Oxford: Young Lives.

Singh, R. and S. Sarkar (2012) *Teaching Quality Counts: How Student Outcomes Relate to Quality of Teaching in Private and Public Schools in India*, Working Paper 91, Oxford: Young Lives.

Tesfay, Nardos K. (2012) 'Horizontal Inequalities in Children's Educational Outcomes in Ethiopia', Student Paper, Oxford: Young Lives.

UNESCO (2015) *Education For All 2000–2015: Achievements and Challenges*, Paris: UNESCO.

UNICEF (2012) *Study on Situation of Out-of-school Children in Ethiopia*, Addis Ababa: ATEM Consultancy Service.

Woldehanna, T. and A. Pankhurst (2014) *Education and Learning: Round 4 Preliminary Findings*, Fact Sheet from the 2013 Young Lives Survey (Round 4) in Ethiopia, Addis Ababa: Young Lives.

Appendix

Table A1. *Descriptive statistics of variables*

	Older Cohort, age 12 (2006 – R2)		Younger Cohort, age 12 (2013 – R4)	
	N	%	N	%
Gender				
Female	479	48.88	884	47.22
Male	501	51.12	988	52.78
Caregiver's education				
None	470	47.96	860	45.87
Adult literacy or religious	153	15.61	164	8.75
Lower primary (1–4)	154	15.71	312	16.64
Upper primary (5–8)	128	13.06	317	16.91
Grade 8 and above	75	7.65	222	11.84
Urban or rural site				
Rural	584	59.59	1,102	58.77
Urban	396	40.41	773	41.23
Region				
Tigray	201	20.51	382	20.42
Amhara	192	19.59	367	19.62
Oromia	200	20.41	392	20.95
SNNPR	244	24.90	463	24.75
Addis Ababa	143	14.59	267	14.27
Type of school				
Private	38	4.09	113	6.39
Community	68	7.33	36	2.04
Government	822	88.58	1,618	91.46
Wealth tercile	328	33.50	630	33.82
Bottom tercile	325	33.20	623	33.44
Middle tercile	326	33.30	610	32.74
Top tercile	328	33.50	630	33.82
Total	**980**	**100.00**	**1,875**	**100.00**

Table A2. *Intake and primary enrolment rates and primary Gender Parity Index by year (%)*

Year	Apparent intake rate	Net intake rate	Gross enrolment rate	Net enrolment rate	Gender Parity Index
2006/7	124	62.6	91.7	79.1	0.87
2007/8	158.4	92	95.6	83.4	0.91
2008/9	162.5	82.2	94.4	83	0.93
2009/10	142.9	76.7	93.4	82.1	0.93
2010/11	163.4	91.3	96.4	85.3	0.94
2011/12	149.3	92.2	95.4	85.4	0.95
2012/13	144.1	95.5	95.3	85.9	0.94

Source: Authors' compilation based on MOE (2006/07) to MOE (2012/3) annual reports.
Note: Apparent intake rate refers to the percentage of new entrants (irrespective of age) in Grade 1 out of the total number of children of the official primary admission age (age 7) in a given year.

Table A3. *Pupil–teacher ratios, and primary school repetition, drop-out and completion rates (%)*

Year	Pupil–teacher ratio	Primary repetition rate	Primary drop-out rate	Primary completion rate
2006/07	59	6.1	12.4	44.9
2007/08	57	6.7	14.6	44.7
2008/09	54	4.9	18.6	43.6
2009/10	51	8.5	13.1	47.8
2010/11	51	8.5	16.3	49.4
2011/12	50	7.9	16.1	52.1
2012/13	50	-	-	52.8

Source: Authors' compilation based on MOE (2006/07) to MOE (2012/3) annual reports.

Table A4. *Maths test score result for OC and YC children (percentage correct scores in maths test) by type of question*

Type of question asked	Percentage correct		
	R2 (2006)	R4 (2013)	Difference (R4–R2)
"2 x 4 ="	83.3	70.8	-12.5
"Which of these is equal to 342?"	64.2	56	-8.2
"Which of these is the name for 9,740?"	75.9	68.6	-7.3
"52-7="	59.4	50	-9.4
"243 + 176 ="	59.3	46.2	-13.1
"It takes Chris 4 minutes to wash a window ..."	54.3	45.7	-8.6
"A piece of rope 204 cm is cut in 4 equal pieces ..."	42.9	27.4	-15.5

Learning Outcomes of Children Aged 12 in Ethiopia:A Comparison of Two Cohorts

The paper examines inequality of learning outcomes among children of the same age (age 12) but seven years apart – in 2006 and 2013. It employs statistical analysis and uses data from the Young Lives Older Cohort who were aged 12 in 2006 (Round 2 of the Young Lives survey) and the Younger Cohort who were aged 12 in 2013 (Round 4), focusing on the Peabody Picture Vocabulary Test (PPVT), mathematics and common reading test results. Changes in test scores for the PPVT, mathematics and reading items were disaggregated by gender, location, region, caregiver education, and whether children attended private or government school. We calculated which groups were showing declining or improving levels of learning more and tried to test the significance statistically. We found that the reading and maths competencies of children had fallen substantially; and girls experienced a higher rate of decline in levels of learning in maths and reading as compared to boys, as did children living in sites in SNNPR and Oromia (compared to children from other regions), rural children (compared to urban children), children in government schools (compared to non-government schools), and children whose caregivers had received little or no education. While the average PPVT raw score of children had increased, rural children and children from SNNPR had seen a decline in their scores; and increased inequality in test results between children in private and government schools and rural and urban children have also been observed. There is an urgent need to halt the increasing learning inequality among children, which may result in inequality in adulthood. The paper proposes further research to identify the in-school and out-of-school factors that contribute to declining levels of learning and widening inequality in learning outcomes.

Young Lives

An International Study of Childhood Poverty

About Young Lives

Young Lives is an international study of childhood poverty, involving 12,000 children in 4 countries over 15 years. It is led by a team in the Department of International Development at the University of Oxford in association with research and policy partners in the 4 study countries: Ethiopia, India, Peru and Vietnam.

Through researching different aspects of children's lives, we seek to improve policies and programmes for children.

Young Lives Partners

Young Lives is coordinated by a small team based at the University of Oxford, led by Professor Jo Boyden.

- *Ethiopian Development Research Institute, Ethiopia*
- *Pankhurst Development Research and Consulting plc, Ethiopia*
- *Centre for Economic and Social Studies, Hyderabad, India*
- *Save the Children India*
- *Sri Padmavathi Mahila Visvavidyalayam (Women's University), Andhra Pradesh, India*
- *Grupo de Análisis para el Desarollo (GRADE), Peru*
- *Instituto de Investigación Nutricional, Peru*
- *Centre for Analysis and Forecasting, Vietnamese Academy of Social Sciences, Vietnam*
- *General Statistics Office, Vietnam*
- *Oxford Department of International Development, University of Oxford, UK*

Contact:
Young Lives
Oxford Department of International Development,
University of Oxford,
3 Mansfield Road,
Oxford OX1 3TB, UK
Tel: +44 (0)1865 281751
Email: younglives@younglives.org.uk
Website: www.younglives.org.uk

Young Lives