Testing Numeric: Evidence from a randomized controlled trial of a computer based mathematics intervention in Cape Town High Schools

1 September 2014

Bianca Böhmer
University of Cape Town

Justine Burns
University of Cape Town

Luke Crowley
Innovations for Poverty Action

Abstract

This paper presents the results of randomized controlled trial conducted to evaluate a Grade 8 after-school mathematics intervention in Cape Town which aimed to fill knowledge gaps using computer assisted learning. The programme employed student coaches to facilitate classes in which Khan Academy resources were used to teach basic numeracy. Large gains of 0.321 standard deviations were observed on basic numeracy outcomes for learners who were selected to be on the programme. Similarly, learners in treatment also scored 0.246 standard deviations higher on core Grade 8 curriculum questions at endline. The improvements in mathematics outcomes were evident for learners throughout the distribution.

Acknowledgements

We would like to thank the Potter Foundation and RUBEN who generously provided the funding for the evaluation. Thanks to Megan Blair for the many helpful conversations. A thank you to the whole research team for the excellent research assistance provided with a particular thanks to Anthea Field, Thomas MBewu, Cecelia Dube, Salma Kagee, William Crawford, Matamela Netshikulwe and Kim Ingle. To the dedicated liaison teachers and administrators at the partner schools as well as the learners who participated in the programme and the evaluation, thank you for the all of the assistance during the year. Many thanks to the Numeric team, particularly the programme managers who helped in numerous ways and Andrew Einhorn who made the intervention possible.

Introduction

Education is viewed as an economic and social priority in almost every country in the world. Human capital formation through education increases the productivity of the labour force and is therefore widely regarded as one of the chief drivers of economic growth (Mankiw, Romer & Weil, 1992, Barro, 2001). Starting in the 1960's, the provision of primary education became an important goal in developing countries. Since then there has been a rapid expansion in access to primary schooling in the developing world (Glewwe & Kremer, 2002).

Despite the successes in increasing school enrollment in developing countries, corresponding increases in learning did not take place. This resulted in a shift away from the primary preoccupation of increasing the quantity of education provided to improving quality. This trend is clearly evident in South Africa where access to education has expanded quickly over the past 30 years at both at the primary and secondary level. Currently, the adjusted net primary enrolment rate[1] is 98.8 percent, and even access to secondary education is high by developing country standards (Statistics South Africa, 2010b, Spaull, 2011). In effect, access to primary schooling in South Africa is universal. However the quality[2] of the education received by the median learner is exceedingly poor.

This lack of quality education can be seen in South Africa's comparatively poor performance relative to developed and middle income countries on international standardised tests. One such example is the 2003 Trends in International Mathematics and Science Study (TIMMS)[3] where South Africa scored last on the list of participating countries (Gonzales et. al., 2004). The provision of education is South Africa is also inefficient as even within the Southern African Development Community (SADC) the Southern and Eastern Africa Consortium for Monitoring Educational Quality (SAQMEC)[4] results show that test scores in South Africa[5] are far below

---

[1] The net primary enrollment rate is defined at the proportion of children aged 7-13 that are attending school.

[2] The term quality of education is used to refer to both the scope and content of the educational material as well as the overall experience and environment. Here the term is used to refer to the learning outcomes of students as measured by standardised tests of particular cognitive skills, such as Mathematics and Science scores on the TIMMS or numeracy and literacy measures of SACMEQ. Such scores are widely available, well defined and easily measurable.

[3] TIMMS is an international comparative study which tests the Mathematics and Science knowledge of learners in the 4th and the 8th grade. It also includes a set of school, educator and learner surveys.

[4] SAQMEC is a set of standardised tests of numeracy and literacy that are conducted with learners in 6th grade in 14 African countries (Spaull, 2011).

[5] South Africa : Percentage of Grade 6 students functionally illiterate, 27.26%, and functionally innumerate, 40.17%,

those of Botswana, a country which spends a similar amount on education per learner (Spaull, 2011).

In South Africa school enrollment is high, thus working to increase the quantity of schooling obtained is futile. Instead, policies or interventions that specifically target learner effort, improve the quality of educational inputs or provide better instruction need to be considered. One possible approach involves introducing new technologies designed to enhance learning, into the homes and classrooms of learners in developing countries.

The advent of computers promised a fundamental change to the organization of education worldwide and Information and Communication Technology (ICT) in the form of computer assisted learning (CAL) holds particular promise for education in developing countries. Although inadequate infrastructure and weak management pose challenges to the introduction of computers in schools, it on of the proposed solution to some of the problems related to teaching and learning in developing countries.

However, simply providing an individual with computers and educational software does not automatically translate into higher scores on cognitive tests of mathematics and language. No effect on mathematics or language test scores was found in Peru, where 902 000 laptops were distributed to school children in rural areas between 2008 and 2009, or in low-income schools in the US, where free computers for home use were randomly assigned to half of the learners in the sample (Crista et. al., 2012, Fairlie & Robinson, 2012). In Romania, experiments showed that a voucher system subsidising home computers had a negative effect on Mathematics, Romanian and English standardised tests scores, with learners in families that received the voucher scoring between 0.25 and 0.33 of a standard deviation lower than learners in control (Malamud & Pop-Eleches, 2011).

Similarly, many state funded programmes, in developed and developing countries, have focused on increasing access to technology in low income schools by providing classroom computers or computer laboratories. This is usually coupled with intensive teacher training and support (Angrist & Lavy, 2002, Louw, Muller & Tredoux, 2008, Barrara-Osorio & Linden, 2009). Ultimately, however, the success of computer assisted learning depends critically on the ability

---

Botswana: Percentage of Grade 6 students functionally illiterate, 10.62%, and functionally innumerate, 22.48%. (Spaull, 2011) of teachers to integrate the new technology into the classroom (Barrara-Osorio & Linden, 2009). The Computers for Education programme in Columbia received refurbished computers from the private sector that were then installed in public schools. A randomized controlled trial of the programme in 97 schools showed that despite receiving training and technical support, survey responses indicated that teachers failed to make use of the computers provided. Thus, there was no significant impact on language outcomes (Barrara-Osorio & Linden, 2009). Even when teachers utilise computers in the classroom this does not guarantee that cognitive outcomes will improve. Angrist and Lavy (2002) investigated the impact of the Tomorrow-98 programme which oversaw the installation of computers in primary and middle schools funded by the Israeli State Lottery. They found that the number of teachers using CAL increased through the programme, but there was no evidence of a simultaneous increase in mathematics and Hebrew test scores of learners in the fourth and eighth grades.

The results of two other randomized controlled trials in which computer aided instruction was used as a supplement in low performance schools were much more positive. In Vadodara in India, Banerjee, Cole, Duflo and Linden (2007) evaluated a programme implemented by an NGO in local schools which used computers provided by the state to set up a weekly class, in which Grade 4 learners played mathematics games on a shared computer for 2 hours a week. The improvement in test scores after a year on the programme stood at 0.35 standard deviations. Additionally, the effect was stronger for children at the bottom of the distribution. A similar finding is one made by Barrow, Markman and Rouse (2008). They looked at a randomized controlled trial in three large underachieving urban educational districts in the US and find that students assigned to classes that were randomly selected to receive instruction on a computer in Pre-Algebra and Algebra scored 0.17 standard deviations higher than learners in the control group.

Additionally, Barrow, Markman and Rouse (2008) note that learners in large classes, classes with a high variation in ability or high absentee rates, benefited more from computer aided instruction in algebra and pre-algebra. This finding is consistent with the hypothesis that CAL enables each learner to advance at their own pace, and is particularly relevant in developing countries, where large class sizes and high absentee rates are real constraints which impede teachers and administrators in providing access to a quality education.

Resource constraints are a real barrier to education in developing countries, however, evidence suggests that a simple transfer or injection of resources is not sufficient to raise the quality of the education received by learners. A wide range of reforms and innovative solutions will be required to extend a quality education to learners in poor districts and remote areas and results from randomized controlled trials suggest that introducing CAL is one approach that has contributed to improving learning outcomes in developed countries.

This paper evaluates Numeric, an organization using computer assisted learning in an after-school programme to improve numeracy levels. A pilot programme was run, in the form of a randomized controlled trial, to evaluate the impact of the programme over the course of one year. The evaluation took place in 9 schools in low income areas in the Cape Town Metropolitan Area, targeting learners in their first year of secondary school. Mathematics tests conducted at endline in November revealed that the programme had a positive effect on mathematics outcomes. The basic numeracy scores of learners that had been selected to participate in the prorgamme were found to be 0.321 standard deviations higher than non-participating learners. Similarly, the test scores of learners in treatment on Grade 8 curriculum material were 0.246 standard deviations above those of learners in the control group.

This paper is divided into eight sections. Section 2 provides a brief overview of the education system in South Africa. Section 3 introduces the Numeric programme and explains features of Khan Academy. Section 4 expounds on details of the experimental design including the school selection procedure as well as the learner application and selection process. Section 5 contains particulars about the data collection process and information about the dataset itself. Section 6 presents the results and analysis of the results. A positive treatment effect, of 0.321 standard deviations on basic numeracy and 0.246 standard deviations on core grade 8 curriculum material is found. Section 7 presents a summary and discussion of the results. Additionally, self-selection during the application process and the after-school nature of the Numeric programme are unpacked and the implication on the evaluation are discussed. Finally, Section 8 concludes.

Education in South Africa

Schooling in South Africa is divided into pre-school, then seven years of primary school, followed by five years of secondary school. The South African Department of Basic Education

(DBE) is responsible for provision of basic education through the provision of learning materials, teacher development, tracking progress through national standardised assessments and development of school management and leadership (Department of Basic Education, 2013a). The DBE is administratively supported by 9 provincial education departments, of which one is the Western Cape Education Department (WCED), which is the province in which the intervention is located.

A number of the challenges facing the South African schooling system are an enduring legacy of the apartheid regime. During this period, the South African education system was segregated along racial lines and non-personnel expenditure was heavily skewed towards white schools and learners (Van der Berg, 2006). Starting in the 1980's changes in policy gradually reduced the spending differentials and the adoption of Norms and Standards in 2000 ratified pro-poor non-personal government expenditure on education. However, despite these shifts in allocation of resources, huge discrepancies in achievement by race and socio-economic status are still evident (Van der Berg, 2007, Spaull, 2011).

Although multiple factors contribute to schools' inability to use this increase in resources to improve educational outcomes, one important factor is a lack of human resources in poorer schools. Good leadership and efficient school management as well as the availability of teachers with adequate content knowledge who take the time to convey this knowledge are paramount to providing a quality education (Taylor, 2011). However, the content knowledge of teachers in South Africa has both a low mean[6] and a high variance (Spaull, 2011, Spaull, 2013). Additionally, the mean level of self-reported absenteeism of South African teachers is high[7]. The combination of low teacher content knowledge and high rates of absenteeism leads to disruptions in teaching and learning and an incomplete coverage of the curriculum. The resulting learning gaps then hamper the future progress of learners.

[6] For example, the mean number of multiple choice questions that Grade 6 teachers answered correctly on the Grade 6 SAQMECIII learner numeracy test was 9 out of 16 items (Spaull, 2013).

[7] Whist teachers in South Africa reported being absent between 11.4 and 23.4 days on average (mean for quintile 5 and mean for quintile 1). Comparatively, the range across the 5 quintiles for Namibia is 9.2 to 9.7 days, in Botswana it is 10.3 to 10.9 days and in Mozambique it is 5.9 to 7.2 days (Spaull, 2011). Although it is possible that some of the difference may be due to a higher rate of systematic underreporting in other countries it is unlikely to account for such a large difference.

The Intervention: Numeric and Khan Academy

Numeric is a non-profit organization that aims to teach basic numeracy and inspire a wider interest in mathematics. Their after-school classes use CAL extensively in the form of free online and offline resources from the Khan Academy website.

Khan Academy

Khan Academy is a non-profit organization, whose mission is summarized as "providing a free world-class education for anyone anywhere" (Khan Academy, 2014). The website consists of three main components: educational videos, the knowledge map and the coach resources. The main method of instruction on Khan Academy is through videos. These enable learners to revise or learn new concepts. The second feature of the website is the knowledge map. The knowledge map is a mathematics exercise framework which allows the learner to practice specific mathematics skills. There are no restrictions governing the order in which exercises need to be completed, so each individual has full autonomy over which exercises they attempt and gamefication is used to incentivize and engage the learners. The third resource that the website provides is a coach tools page which provides detailed statistics on each learner's progress.

Numeric

The Numeric after-school intervention was the programme under evaluation. There were a total of 11 classes in the pilot from 9 schools in the Cape Town Metropolitan Area.

The biggest technical innovation was the creation of the offline video browser (OVB). Due to bandwidth constraints in South Africa it is infeasible to stream videos over the Internet in a classroom environment. Instead, all of the videos that are on the Khan Academy website were centrally downloaded and are then stored locally on a server or hard drive. Users are then able to watch the videos by streaming it from the server.

The programme took place after-school during the week or on a Saturday morning. The learners attended Numeric classes for an hour and a half on a bi-weekly basis. The school computer laboratory was used as the venue for the sessions. The coaches, generally mathematics teachers-in-training, were hired and trained by Numeric at the beginning of the year.

The learners were encouraged to watch a video at least once per session, but the majority of class time was spent doing mathematics exercises online using the Khan Academy Knowledge Map. However arithmetic drills and mathematics games were also incorporated into the classes.

In contrast with most school related mathematics interventions the Numeric programme is not curriculum based. They do not align their schedule with the schools schedule and at no point during the year do they deliberately set aside time to tutor or revise sections covered in class. Instead they start out the year with basic arithmetic and focus on basic numeracy.

Parental involvement was encouraged through a parent's evening and by sending home reports. The learners were also required to pay a nominal participation fee on a quarterly basis, however a fee waiver option was available to all applicants, to avoid excluding learners on financial grounds. The programme had an attendance requirement of 75 percent that was stringently enforced.

The Numeric programme is a combination of all of the features mentioned. Although using computer assisted learning as a tool to teach and practice mathematics is the core objective of Numeric, any effect needs to be viewed as an outcome of the programme as a whole.

Experimental Context

The Numeric intervention requires access to a networked computer laboratory, with a connection of at least 4 MB/s. This is a stringent requirement in a country where access to computers is still low and broadband is still slow and prohibitively expensive for the majority of the population (World Bank, 2014, Mybroadband, 2014).

Access to computers at schools is slowly improving. In 2002 39.2 percent of all schools in South Africa had computers, although only 26.5 percent were for teaching and learning (Department of Basic Education, 2004). Whereas in the 2009 Census at School survey the percentage of learners with access to a computer at school was 53 percent, however, only 20 percent of learners in Grade 3 to 12 reported that they had access to the Internet (Statistics South Africa, 2010a). The 2004 White Paper on e-Education stated that computer access and internet connectivity are necessities at both primary and secondary level, however there has been no coordinated national programme. This lack of centralisation has resulted in huge variation in availability of computers and access to the internet across provinces (Department of Basic Education, 2004).

In the Western Cape a programme named Khanya was spearheaded by the Western Cape Education Department between 2002 and 2012. Through Khanya every secondary school in the province was equipped with a computer laboratory by 2006 (Western Cape Education Department, 2009). Thus, even in poorer areas, some schools still had the required infrastructure in place in 2012 or could easily upgrade their facilities to meet the requirements for the Numeric programme. This made the Western Cape an ideal environment to test the programme without incurring the initial expense of setting up computer laboratories with a local area network.

Experimental Design

The evaluation was designed to take the form of a randomized controlled trial[8], with the unit of randomisation at the individual level. The selection of participants took place in two stages. In the first stage schools applied and were selected. This was followed by a set of baseline tests and surveys which were conducted in the selected schools. Thereafter, a learner application and selection process took place. Applicants at each school were either selected to be in the treatment or control group, or were placed on a waiting list.

School Selection

The schools were selected by the implementing organization in collaboration with the circuit managers of the Western Cape Department of Education (WCED) given the following two criteria for participation: good management and a working computer laboratory with an internet connection.

Numeric met with the selected schools, explained the programme and handed out application forms. Once the school applications had been received, the Numeric technical team conducted a second thorough inspection of the computer labs and internet connectivity at each school. Updates to their system or internet connectivity speeds were recommended as required.

[8] A randomized controlled trial (RCT) was chosen as the method of evaluation, as selection into such a programme is a determinant of observable and unobservable learner characteristics. It would therefore be unwise simply to compare the outcomes of learners on the programme and learners that chose not to participate, without correcting for this selection bias. Although econometric methods such a two stage least square estimation or matching can be used for such corrections, these often require very large and diverse samples and the existence of a valid instrumental variable. Randomized controlled trials instead allow for the calculation of an unbiased treatment effect through the creation of a control and treatment group which are comparable- on average- at baseline. An average treatment effect can then easily be calculated, without further corrections, which is free of selection bias despite the presence of unobservables.

Implications for the External Validity of the programme

It is important to note at this point that the school selection process has definite implications for the external validity of the evaluation. The Western Cape province is not representative of the country as a whole. The Western Cape is a relatively wealthy province with a high proportion of urban households (Finn, Leibbrandt & Woolard, 2009). It has also consistently been one of the two provinces achieving the highest pass rates in the annual National Senior Certificate Examinations[9]. There were also three distinguishing characteristics of selected schools that differentiated them from a representative sample within the Western Cape. The schools were urban, all were well managed with strong leaders and they had functioning computer laboratories with a working internet connection.

The fact that a school was required to have a functional computer laboratory does not have as great a repercussion for the external validity of the programme. This is because the Khanya project was an external shock, implemented by the Western Cape Education Department. Thus due to this government intervention the socio-economic status of leaners at a school is not necessarily an indicator of the digital resources available to the school. Consequently, the third characteristic of a functional computer laboratory would be correlated with better management insofar as the computers would need to be serviced and maintained, which is more likely to take place in a well-managed school. However this selection criterion is not necessarily a good gauge of other school or learner characteristics.

Learner Application and Selection

The Numeric programme is an optional after-school activity. Additionally, due to resource constraints there were a limited number of places available on the programme. As a result, learners had to apply to be considered. From this group of applicants, learners were then randomly selected to be on the programme or in the control group, whilst the remainder was placed on a waiting list. If one of the treatment learners dropped out of the programme, then a learner from the waiting list could then be offered their place on the programme. Table 1 shows

[9] For the first time in a number of years, neither Gauteng nor the Western Cape had the highest pass rates in January 2014. Instead they were ranked 3rd and 4th respectively, whilst the Free State and Northern Cape were place 1st and 2nd, sparking debate around learner dropout which questioned the validity of these figures (John, 2014)

the total number of learners in Grade 8 at each school[10]. Column 2 of Table 1 shows the number of applications received at each school.

Table 1: Table to show the number of learners, number of applicants and acceptance status at each of the nine schools

School	Number of Learners in Grade 8	Applications Received	Number of Learners in the Evaluated Sample
101	197	43	40
102	161	45	40
104	183	51	40
105	272	101	80
106	266	83	80
107	285	65	40
109	230	34	32
110	170	53	40
111	244	93	80
Total	**2008**	**568**	**472**

Source: Own data (2013)

Successful applicants received a letter of acceptance a few days prior to the first Numeric class at their school. The starting dates of the nine schools were staggered over five weeks.

Applicant Characteristics

Table 2 presents some of the differences in observable characteristics of applicants. The first row shows that the average numeracy score of applicants is a significant 2.2 percentage points higher than non-applicants.

The gender of the learners also plays an important role in the decision to apply which can be seen by the large number of female applicants in row 2 of Table 2. Whereas 60 percent of applicants are female, only 52 percent of learners in Grade 8 at the nine participating schools where female. The gender difference between the applicant and the non-applicant group is highly significant.

The age difference between applicants and non-applicants is also highly significant, with the average age of applicants 2.4 months lower than that of non-applicants (Table 2, Row 3). Part of the explanation for this difference is due to grade repetition. It is evident that a non-applicant is 15.5 percent more likely to have repeated a grade than an applicant. However it is likely that

[10] As many of the schools did not have reliable learner numbers at the beginning of the year, class registers from the end of the second term were used to calculate the learner numbers in the first column of Table 1.

there are factors other than grade repetition, both observable and unobservable, that contribute to this difference in the mean ages.

Table 2: Table showing baseline characteristics of applicants and non-applicants

VARIABLES	Non-Applicant	Applicant	**Difference** *(Applicant-Non-Applicant)*
Baseline – Mathematics Section A Test Scores	0.435 (0.179)	0.457 (0.165)	0.022** (0.009)
Female	0.485 (0.500)	0.605 (0.489)	0.120*** (0.024)
Age	14.347 (0.992)	14.148 (0.872)	-0.199*** (0.046)
African	0.447 (0.497)	0.452 (0.498)	0.005 (0.025)
Asian, Indian or White	0.033 (0.178)	0.020 (0.139)	-0.013* (0.008)
Coloured	0.521 (0.500)	0.529 (0.500)	0.008 (0.025)
Home Language: Afrikaans	0.137 (0.344)	0.091 (0.288)	-0.047*** (0.015)
Home Language: English	0.448 (0.497)	0.473 (0.500)	0.025 (0.025)
Home Language: IsiXhosa	0.403 (0.491)	0.425 (0.495)	0.022 (0.025)
Home Language: Other Language	0.012 (0.107)	0.011 (0.103)	-0.001 (0.005)
Learner's Mother is Alive	0.961 (0.193)	0.959 (0.198)	-0.002 (0.01)
Learner's Mother has completed Matric	0.571 (0.508)	0.568 (0.504)	-0.003 (0.028)
Learner's Father has completed Matric	0.601 (0.504)	0.608 (0.499)	0.007 (0.03)
Travel Time to School	22.097 (16.514)	21.839 (15.995)	-0.258 (0.822)
At least one computer at Home	0.672 (0.470)	0.607 (0.489)	-0.064** (0.026)
Internet on a Computer at Home	0.404 (0.491)	0.358 (0.480)	-0.046* (0.028)
Learner has repeated at least one Grade	0.319 (0.466)	0.277 (0.448)	-0.042* (0.024)
Like Mathematics A Lot	0.389 (0.488)	0.461 (0.499)	0.072*** (0.026)
Satisfaction (Scale 1-7)	5.333 (1.644)	5.428 (1.560)	0.095 (0.083)
Has a Phone	0.813 (0.390)	0.783 (0.412)	-0.030 (0.021)
Household members	5.344 (2.187)	5.297 (2.496)	-0.046 (0.123)
Number of Observations	**1543**	**569**	**2112**

Notes: Column 1 and column 2 have standard deviations in parentheses. Standard errors are in parentheses in last column. Statistically significance differences are indicated by: *** p<0.01, ** p<0.05, * p<0.1

Data

Data Collection: Test and Survey Schedule

Data collection took place during three distinct periods over the course of the year. Baseline tests and surveys were administered in January and February and, for one of the schools, a second round of tests was held in the first week of March[11]. All baseline testing was completed prior to randomisation. Midline testing was completed in June and endline surveys and tests were held in November.

A complete schedule of the tests and surveys is outlined in Table 3. Row 1 and 2 show that three standardised mathematics tests were conducted over the course of the year. In order to simplify comparisons across time, the format of the test and question structures were identical in all three testing periods, although the numbers[12] were altered for each test. The full test consisted of two sections: Section A and Section B. Section A covered primary school curriculum, whilst Section B tested concepts taught in Grade 7 and 8. The demographics surveys were also administered to all Grade 8 learners at baseline and endline.

Table 3: Table indicating during which testing periods particular tests and surveys were conducted and which group of learners were tested or surveyed

Test or Survey description	Baseline	Midpoint	Endline
1. Standardised Mathematics Test – Section A (Numeracy)	A	A	A
2. Standardised Mathematics Test – Section B (Core Grade 8 curriculum)		A	A
3. Demographics and Access to Technology Survey	A		A

Note: The letters indicate which group of learners wrote the test or completed the survey. A - All Grade 8's at participating schools, P – all applicants to the programme, E - Learner in the evaluated sample (control and treatment).

Dataset

The complete dataset contains observations on 2112 Grade 8 learners, originating from the nine schools. Any leaner that wrote at least one of the tests or surveys was assigned a unique number

[11] The set of tests that were administered in March consisted of the Raven's Progressive Matrices, a test of English Literacy, and a Computer Literacy and typing test.
[12] New numbers were used in each testing period, but the question was of the same type. For example Question 1.1 is 5×6 at baseline, 4×7 in the midline test and 6×7 at endline.

and was added to the dataset. Therefore, this total includes all Grade 8 learners from all nine schools, even those that moved schools during the year. It comprises of the results of the three mathematics tests, as well as the responses on the baseline and endline learner surveys.

As the primary focus of this paper is to determine the magnitude of the treatment effect on mathematics outcomes, the main dataset that is utilised is a subset of the full dataset. This dataset contains only the 472 learners who were selected to be in the treatment or the control group. This subset is referred to as the evaluated sample, and consists of 236 individuals who comprise the control group and 236 learners in treatment. .

Missing Observations and Learner Attrition

Table 4 shows the proportion of observations obtained from the evaluated sample for each test and survey. Rows 1 to 3 show that there was a fairly high rate of test completion on the main measure, the standardised mathematics test. Of the sample of 472 learners, 97.5 percent of the sample wrote the baseline mathematics test, whilst 96.0 percent wrote the endline mathematics test. The proportion of completed endline surveys was slightly lower compared to the proportion of mathematics test written at endline, with a coverage of 94.9 percent. There was no difference in response rates between control and treatment on the endline survey. The number of missing observations for the baseline survey was very small at only 2.54 percent.

Table 4: Table to show the proportion of the evaluated sample that wrote the tests and completed surveys for each testing period by treatment and control

	Control	Treatment	Total	Significant Difference *(Treatment-Control)*
Baseline - Mathematics Test	0.979	0.975	0.977	-
Midline - Mathematics Test	0.919	0.941	0.930	-
Endline - Mathematics Test	0.962	0.958	0.960	-
Baseline - Survey	0.987	0.962	0.975	*
Endline - Survey	0.949	0.949	0.949	-

Notes: Numbers in Columns 1, 2 and 3 are means. Statistically significance differences are indicated by, *** p<0.01, ** p<0.05, * p<0.1, in the 4th column.

Analysis

Summary Statistics

The credibility of the results in a randomized controlled trial depends critically on the random selection into treatment and control at baseline. Table 5 shows the means and the differences in the means of the control and treatment groups at baseline. Only one of the variables in question was found to be significantly different in control and treatment, using a basic t-test with unequal variance. The variable is a combination of all learners who are Asian, Indian or White. There are only 10 individuals or 2.15 percent of sample learners who fall into this category, thus the small size of this group is the probable cause of the imbalance between treatment and control.

It is interesting to note that a substantial proportion of the sample, 62.9 percent, has access to at least one computer at home. Internet access is less common. Only 37 percent of learners report having internet access on a computer at home. This shows that exposure to technology is already reasonably high prior to the start of the programme.

Table 5: Baseline learner characteristics in the control and treatment groups

VARIABLES	Control	Treatment	Difference *Treatment-Control*
Baseline-Mathematics Test Scores (Section A)	0.464	0.463	-0.001
	(0.165)	(0.171)	(0.016)
Baseline -English Literacy Score	0.558	0.563	0.243
	(0.197)	(0.186)	(0.020)
Baseline – Raven's Score	0.593	0.576	0.017
	(0.169)	(0.175)	(0.017)
Baseline -Typing (Keys per minute)	0.608	0.602	-0.006
	(0.330)	(0.336)	(3.808)
Female	0.555	0.623	0.068
	(0.498)	(0.486)	(0.045)
Age	14.106	14.177	0.07
	(0.753)	(0.946)	(0.079)
African	0.391	0.438	0.047
	(0.489)	(0.497)	(0.046)
Asian, Indian or White	0.034	0.009	-0.026*
	(0.182)	(0.092)	(0.013)
Coloured	0.575	0.554	-0.021
	(0.495)	(0.498)	(0.046)
Home Language: Afrikaans	0.086	0.103	0.017
	(0.281)	(0.304)	(0.027)
Home Language: English	0.532	0.479	-0.053
	(0.500)	(0.501)	(0.046)
Home Language: IsiXhosa	0.373	0.402	0.028
	(0.485)	(0.491)	(0.045)
Home Language: Other Language	0.009	0.017	0.009
	(0.092)	(0.130)	(0.010)
Learner's Mother is Alive	0.953	0.965	0.012
	(0.213)	(0.185)	(0.019)
Learner's Mother has completed Matric	0.585	0.545	-0.040
	(0.504)	(0.509)	(0.051)
Learner's Father is Alive	0.860	0.859	-0.001
	(0.347)	(0.349)	(0.033)
Learner's Father has completed Matric	0.590	0.653	0.063
	(0.505)	(0.490)	(0.054)
Number of Household Members	5.350	5.413	0.063
	(2.164)	(3.056)	(0.248)
Travel Time to School	19.382	21.231	1.849
	(12.735)	(15.523)	(1.125)
At least one computer at Home	0.614	0.640	0.026
	(0.488)	(0.481)	(0.047)
Internet on a Computer at Home	0.335	0.410	0.075
	(0.473)	(0.493)	(0.051)
Learner has repeated at least one Grade	0.257	0.308	0.051
	(0.438)	(0.463)	(0.042)
Learner Likes Mathematics a Lot	0.453	0.467	0.014
	(0.499)	(0.500)	(0.047)
Number of Observations	**236**	**236**	**472**

Notes: Column 1 and column 2 have standard deviation in parentheses. Standard errors are in parentheses in last column. Statistically significance differences are indicated by, *** p<0.01, ** p<0.05, * p<0.1, in the third column

Treatment Effect Size

The main objective of the Numeric after-school programme was to improve the mathematics outcomes of participants. As such, the primary focus of the analysis was on determining the overall effect of the programme on the mathematics scores of learners, and then checking the consistency and robustness of the results.

First, a basic intention-to-treat (ITT) approach was used to calculate the effect of treatment. Under ITT learners that were selected to be in treatment, regardless of actual treatment status, counted as treated learners. The same applies for the control group.

In this experiment, there was a high correlation between selection into treatment and actual treatment status. None of the control learners were ever treated. In comparison, 97 percent of learners in treatment attended at least one class, whilst 71 percent of learners received a full year of treatment.

The effect size was determined using baseline and endline mathematics test scores and can be calculated using Glass's delta (δ). Glass's delta is defined as the difference in the means of the treatment and control groups divided by the standard deviation of the control group (Glass, 1977).

$$\delta = \frac{\mu_{treatment} - \mu_{control}}{\sigma_{control}}$$

The measure δ presents the effect size in units of standard deviations (s.d.). This allows for greater comparability across different tests and studies, by providing a measure of the magnitude of the difference in outcomes between treatment and control.

In this paper, only baseline and endline scores are used in the calculation of effect size. The midline tests are used to determine learning trajectories, but are not included in the main analysis. An approach using only baseline and endline measures is likely to produce results with a large error if the outcomes are noisy. Here, however, the endline test results on Section A and B are highly correlated with other measures of mathematics aptitude[13]. This suggests that they provide a fairly accurate measure of overall mathematics ability at endline, allowing for slight

[13] The correlation between the incentivised test marks and the endline test Section A and Section B marks are high at 0.878 and 0.819 respectively. Likewise the correlation between second term school marks and endline Section A results is also high at 0.606, whilst it is 0.678 for endline Section B scores. This last result is unexpectedly high, considering that this does not account for differences in reporting between schools and teachers.

differences in performance of the learner on the days of the test, different testing conditions and differences between markers[14]. Due to the relatively small error in the outcome measure and high correlation between baseline and endline mathematics scores, this approach is justified (McKenzie, 2012).

This section comprises of two parts and presents six results. The first subsection examines, in detail, the trends and distributions in average mathematics outcomes. The effect size of the Numeric programme is also calculated. Six major results are discussed in this subsection. Namely, (1) that there is an observable programme effect at midline which becomes highly significant at endline. The results on the endline mathematics test are also found to be significantly higher for treatment learners in (2) basic numeracy and for (3) core Grade 8 curriculum. Additionally, these (4) improvements are evident for learners throughout the distribution. Furthermore, (5) learners who were selected to be in treatment are significantly more likely to state that they like mathematics than control group learners at endline.

The second part takes a closer look at attendance and drop-out rates and how these are related to the programme effect and finds that (6) Numeric drop-out rates are negatively related to baseline mathematics scores.

Mathematics Test Results

Trends in the Mathematics Test Scores

Result 1: A continuous improvement in average mathematics scores of treatment learners, relative to control learners, is observed over the course of the year.

Table 6 reports the average scores on the mathematics tests over the three testing periods. Columns 1 and 2 show the means of the control and treatment groups, whilst Column 3 reports the difference between treatment and control. Row 1 shows that at baseline, there is no significant difference in the means, with mean scores of 46.4 and 46.3 percent respectively for the control and treatment groups. By midline, differences start to emerge; the treatment group outperforms the control group by about 3.5 percentage points on the basic numeracy test (Section A), significant at a 5 percent level, whereas the treatment group scores an insignificant 1.3

[14] The error rate due to marking and capturing was estimated to be 1.0 percent.

percentage points higher on core Grade 8 curriculum material (Section B). The treatment group scores significantly higher in both sections at endline, with a difference of 6.0 percentage points (0.321 s.d.) in Section A and an average of 4.1 percentage points (0.246 s.d.) in Section B, significant at a 1and 5 percent level of significance.

Since the three basic numeracy tests were of a comparable standard, the Section A results can be compared across time. Column 1 of Table 6, shows that even in the absence of any outside intervention the average score of the control group increased by 2.4 percentage points from baseline to endline[15]. Comparatively, in the treatment group, there is an even larger increase of 5.2 percentage points in basic numeracy scores from baseline to endline.

Table 6: Mathematics test scores at baseline, midline and endline for treatment and control

VARIABLES	Control	Treatment	Difference *Treatment-Control*
Baseline - Mathematics Section A	0.464	0.463	-0.001
	(0.165)	(0.171)	(0.016)
Midline - Mathematics Section A	0.512	0.548	0.035**
	(0.172)	(0.191)	(0.017)
Midline - Mathematics Section B	0.241	0.254	0.013
	(0.143)	(0.163)	(0.015)
Endline - Mathematics Section A	0.488	0.548	0.060***
	(0.187)	(0.196)	(0.018)
Endline - Mathematics Section B	0.264	0.305	0.041**
	(0.167)	(0.194)	(0.017)

Notes: Column 1 and column 2 have standard deviations in parentheses. Standard errors are in parentheses in last column. Statistical significance is indicated by *** p<0.01, ** p<0.05, * p<0.1, in the third column.

Multivariate Analysis

The main analysis investigates the effect of selection into the programme on the mathematics scores of learners at endline. Here the equation of interest is:

$$Y_{isEL} = \propto + \beta T_{is} + \gamma Y_{isBL} + \pi Y_{isBL}^{2} + \theta \boldsymbol{S'}_{s} + \rho \boldsymbol{Z'}_{isBL} + \delta \boldsymbol{X'}_{is} + \varepsilon_{is} \qquad (1)$$

[15] Looking only at the control group it is evident that the Section A results increase significantly from baseline to midline and then drop by about 2.5 percent between midline and endline. There are a number of reasons that may have contributed to this pattern. Section A does not test Grade 8 material, so whilst the curriculum provides for revision at the beginning of the year, learners would not normally encounter these types of problems in their school Maths classes in the third and fourth terms, as a result they may forget some of the methods. The marking at endline was also stricter for the longer questions, which potentially reduced marks by a further percent or two. Also, learners were only given feedback on their performance on these tests at the very end of the year and this lack of feedback may have resulted in reduced effort on the endline test.

The dependent variable, Y_{isEL}, is the mathematics test score of learner i at school s at endline (EL), whist Y_{isBL} is their baseline (BL) mathematics score. A second order term, πY_{isBL}^2, was included to allow for a non-linear relationship between baseline and endline mathematics scores. T_{is} is a dummy variable which indicates whether the learner was selected to be in treatment and $\boldsymbol{S'}_s$ is a vector of school dummies.

$\boldsymbol{Z'}$ is a matix of tests scores at baseline which include English literacy, the Raven's Standard Progressive Matrices results as well as a typing speed test. As the number of missing variables at baseline was reasonably high, and the missing values were not missing completely at random, a missing indicator method approach was adopted (White & Thompson, 2005)[16]. Namely, the missing values on all covariates were set to zero. Furthermore, a dummy variable was included for each covariate which was equal to 1 if the covariate had missing values and zero for all non-missing observations.

$\boldsymbol{X'}$ is a matrix of learner characteristics. These include demographic details such as age, gender and home language. Variables measuring attitudes towards school, access to technology, parent's education and some household related statistics were also included in $\boldsymbol{X'}$. Lastly, ε_{is} is the error term for individual i.

Result 2: Treatment learners score significantly higher on basic numeracy at endline than control learners.

Table 7 presents the main regression results. In the first four columns the dependent variable is the Section A mathematics test score at endline, whilst Columns 5 to 8 have the endline Section B results as the regressand. The first column shows a treatment effect of 6.0 percentage points with no controls. In Column 2 a linear and second order term of baseline mathematics results were included along with a set of school dummies to control for school level differences. This improves the fit of the model, but does not change the significance of the treatment effect whereas the size decreases marginally to 5.9 percentage points. It is also evident that baseline Section A scores are a good predictor of endline Section A scores. Although provision is made for a non-linear relationship between baseline and endline Section A mathematics results the

[16] Both missing observations and item non-response were not random. More saliently, they were not uncorrelated with the outcome variable, namely endline Mathematics results.

relationship is found to be linear, with a highly significant estimate on the linear term, whilst the estimate on the second order term is practically small and statistically insignificant.

In Column 3 there is no material change to the size or significance of the treatment effect when learner demographic controls and the baseline test results on the Raven's, English Literacy and typing tests are included, nor is there a change to the treatment effect when the full set of controls are included in Column 4. As there appears to be no significant change due to the addition of the controls in Columns 4 and 8, the specification which is used repeatedly for the remainder of the thesis is the simpler functional form presented in Columns 3 and 7, with minor alterations where necessary.

The endline basic numeracy results consistently show a difference of between 5.7 and 6 percentage points of treatment learners relative to control group learners. Given a standard deviation of 18.7 percentage points in the basic numeracy scores at endline, this equates to an improvement of 0.321 of a standard deviation in basic numeracy. Alternatively, an average learner who was selected to be in treatment is able to answer 12 percent more of the paper correctly than a learner from the control group, holding individual, school and household characteristics as well as baseline test scores constant. This result is large relative to the effect sizes that are typically observed in education interventions[17]. However, it is of a similar magnitude to the highly successful primary school computer assisted learning programme run in Vadodara, India. This programme also had a mathematics focus and the treated learners improved by 0.35 standard deviations more than control learners in the first year of the programme (Banerjee et. al., 2007).

Result 3: Treatment learners score significantly higher than control learners on core Grade 8 curriculum at endline.

Section B tests Grade 8 curriculum material, which is never explicitly taught on the Numeric programme. However, it is assumed that a better understanding of underlying mathematical principles as well as an increase in the time spent practicing mathematics would contribute to an improvement in school Mathematics grades. The estimates of treatment on Section B scores are

[17] Kremer, Brannen & Glennerster (2013) present a summary of the results of 28 different randomised controlled trails on primary school education interventions. The interventions differ widely both in locality and type of intervention, but out of the 28, only 7 resulted in an improvement in tests scores of more than 0.2 standard deviations.

reported in Columns 5 to 8. They mirror the results in Columns 1 to 4. The improvement due to treatment on core grade 8 curriculum is smaller than the improvement on basic numeracy, but still highly significant.

The average treatment effect is 4.1 percentage points in Column 5. This is equivalent to an effect size of 0.234 standard deviations. In Column 6, when baseline mathematics results and school effects are held constant, the treatment effect falls slightly to 3.7 percentage points. In column 7 individual characteristics and baseline tests scores on the Raven's, English Literacy and typing test are added as controls, whilst in Column 8 the full set of controls is added. The coefficient on treatment is significant at the one percent level in all regressions. This shows that learners in treatment are achieving results that are between 3.7 to 3.9 percentage points higher than control group learners, holding individual demographic and household variables, school and baseline test results constant. The mean score of the control group was 26.4 percent on core Grade 8 curriculum at endline. This means that holding baseline mathematics results, school, individual and household characteristics, baseline scores on the English, Raven's, and typing tests constant, the average treatment learner was able to answer 15 percent more questions correctly than a comparable control group learner.

Table 7: OLS Estimates of treatment on mathematics scores on Section A and Section B at endline

	Endline Mathematics Test Scores							
	Basic Numeracy - Section A				Grade 8 - Section B			
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	0.060***	0.057***	0.058***	0.058***	0.041**	0.037***	0.039***	0.037***
	(0.018)	(0.010)	(0.010)	(0.010)	(0.017)	(0.011)	(0.010)	(0.011)
Baseline-Section A		0.845***	0.716***	0.725***		0.175	0.025	0.061
		(0.133)	(0.140)	(0.151)		(0.132)	(0.130)	(0.139)
$(\text{Baseline-Section A})^2$		-0.000	0.020	-0.012		0.619***	0.641***	0.591***
		(0.126)	(0.129)	(0.138)		(0.131)	(0.124)	(0.133)
Female			-0.022**	-0.008			-0.003	0.005
			(0.010)	(0.012)			(0.012)	(0.013)
Age			0.334***	0.291**			0.306***	0.377***
			(0.105)	(0.118)			(0.104)	(0.116)
Baseline – Raven's Score			0.107***	0.093**			0.155***	0.131***
			(0.035)	(0.039)			(0.035)	(0.037)
Baseline -English Literacy Score			0.076**	0.095**			0.112***	0.129***
			(0.038)	(0.039)			(0.037)	(0.039)
Baseline -Typing (Keys per minute/100)			0.029	0.026			0.020	0.026
			(0.021)	(0.021)			(0.023)	(0.023)
At least one computer at Home				-0.011				-0.006
				(0.015)				(0.014)
Internet on a Computer at Home				0.008				-0.007
				(0.013)				(0.014)
Travel Time to School				0.001				-0.000
				(0.028)				(0.033)
Learner Likes Mathematics a Lot				0.024**				0.020
				(0.011)				(0.012)
Constant	0.488***	0.063*	0.438***	0.384***	0.264***	-0.041	0.163	0.209
	(0.012)	(0.037)	(0.129)	(0.137)	(0.011)	(0.037)	(0.126)	(0.140)
School Controls	N	Y	Y	Y	N	Y	Y	Y
Learner Demographic Controls	N	N	Y	Y	N	N	Y	Y
RCE Controls	N	N	Y	Y	N	N	Y	Y
Household, education and attitude Controls	N	N	N	Y	N	N	N	Y
Observations	453	453	453	453	453	453	453	453
R-squared	0.024	0.707	0.741	0.754	0.013	0.627	0.668	0.687

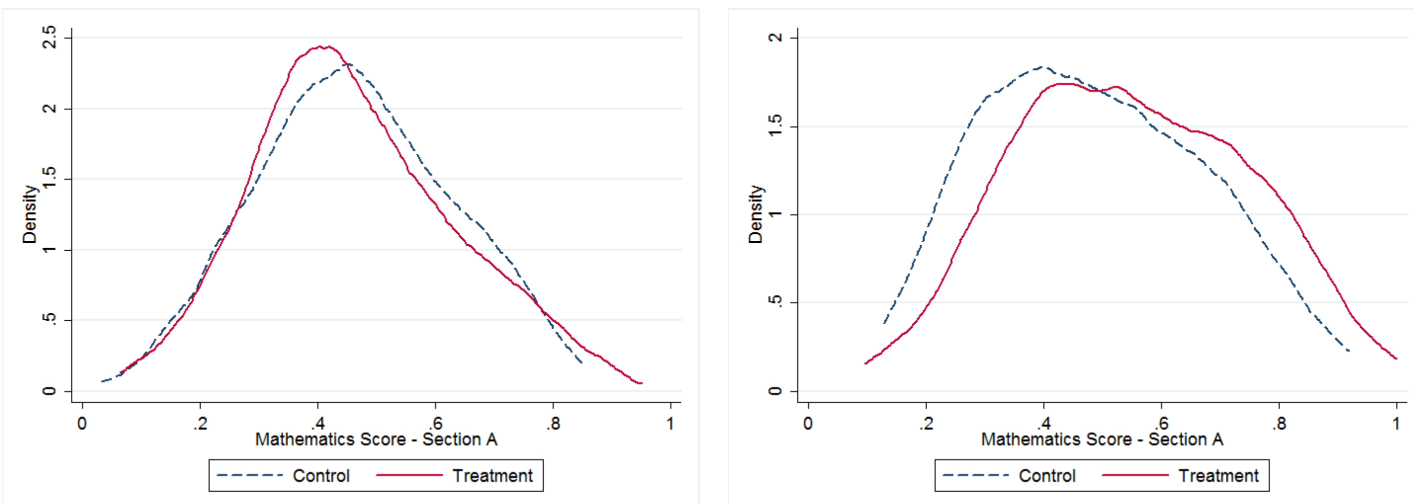
Notes: Robust standard errors are in parentheses. Statistical significance is given by *** p<0.01, ** p<0.05, * p<0.1. School controls consist of a dummy variable for each school. Learner demographic controls include race and home language. RCE controls also include a dummy variable indicating if there was a missing observation for the English literacy, Raven's or typing test at baseline. Household, education and attitude controls consist of mother alive, father alive, mother has matric, father has matric, learner has repeated at least one grade, participant in a sports team or cultural group as well as dummy variables for each covariate that is equal to 1 if the item for that individual is missing.

Result 4: A significant improvement in basic numeracy and core Grade 8 curriculum is observed for treatment learners across the full distribution.

Even with the sparse information provided in Table 6, it is clear that the standard deviation (17.2 percent) of the treatment group at baseline is higher than the standard deviation (16.5 percent) in the control group. This difference is evident in the distribution observed in Figure 1a below. At baseline, there are marginally more learners in the treatment group that are obtaining marks below 20 percent. However, the two distributions are reasonably similar for learners scoring below 25 percent. The difference is evident in the middle and upper ends of the distributions. The control plot lies to the right of the treatment plot for all values between 30 and 80 percent. This shows that the bulk of the control group is receiving higher scores compared to equivalently ranked individuals in treatment. However, of the learners scoring above 80 percent, the majority are in treatment. As a number of the top achieving learners in the sample are in treatment, the analysis of the distribution as a whole becomes particularly important.

Figure 1: Distribution of mathematics scores – Basic numeracy (Section A) at baseline and endline

a. Baseline **b. Endline**



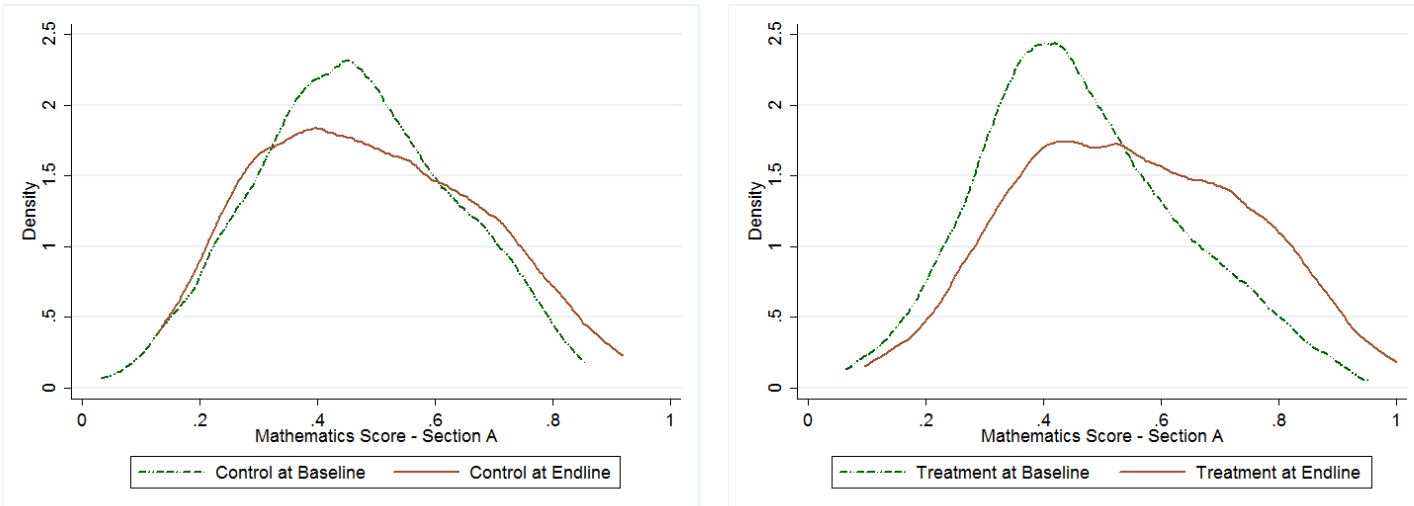
Source: Own data collected in 2013

Figure 1b shows the endline distribution of basic numeracy scores of the treatment group relative to the control group. It is evident that with the exception of a few of learners at the very bottom of the distribution, the treatment function lies to the right of the control function at every point along the distribution.

Figure 2 compares the baseline and endline distributions within the control and treatment groups. The overall increase in the mean of the combined treatment and control groups is 5.16 percentage points. There is, however, also an increase in the standard deviation in both the control and treatment group’s endline numeracy scores. The first graph of Figure 2 shows that from baseline to endline the distribution of numeracy scores in the control group flattens and moves marginally to the right. At endline, there is an increase in the number of learners scoring above 60 percent. However, whilst it is clear that many of the learners are performing better, the proportion of learners achieving less than 30 percent has also increased. In comparison in Figure 2b it is evident that in the treatment group the entire distribution has shifted to the right by endline. Here it can be observed that even the very lowest performing students in treatment have improved over the course of the year. Results from prior studies suggest that learners at the bottom of the distribution benefit the most from computer assisted learning (Banerjee et. al, 2007, Barrow, Markman & Rouse, 2009). Here, however, it appears that learners throughout the distribution are improving, potentially, by similar amounts. Estimates reported in Table 19 serve to confirm this hypothesis.

Figure 2: Distribution of mathematics scores – Basic numeracy of treatment and control groups comparing baseline and endline

a. Control **b. Treatment**



Source: Own data collected in 2013

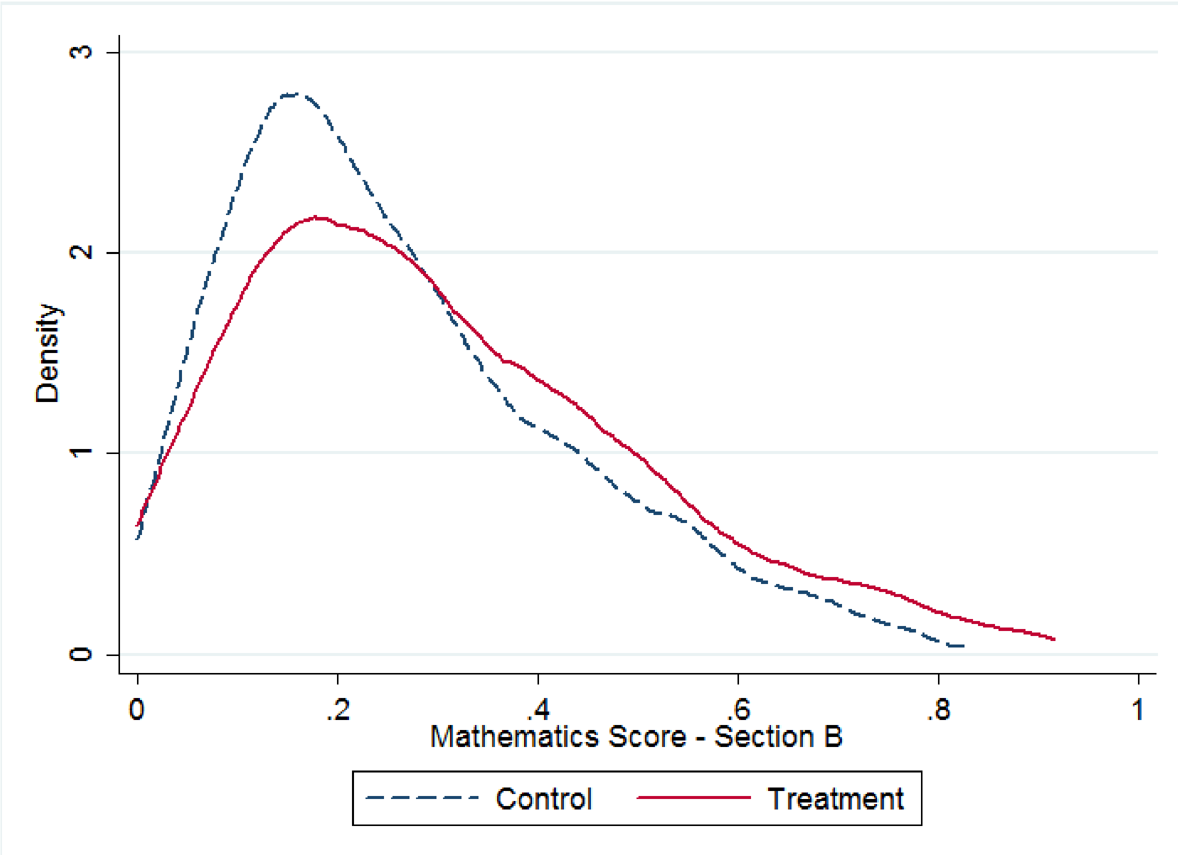
Figure 3 shows the distribution of the endline Section B mathematics scores of the control and the treatment groups. The average score obtained for Section B was considerably lower than the average score achieved in Section A. The scores are generally very poor with 75 percent of learners achieving below 40 percent at endline on core Grade 8 material. Also, whilst the distribution of the Section A scores approximated those of a normal distribution, the distribution of the Section B scores more closely resembles a log-normal distribution with a median at around 20 percent and a long tail to the right,

The general pattern, when comparing the control and treatment groups, is similar to the one observed in the endline Section A distributions presented in Figure 1b. This same pattern is observed in Figure 3 where it is evident that the treatment scores are higher than the scores obtained by control group learners at all points along the distribution, with the exception of a small proportion of learners at the very bottom. In contrast to the Section A results where the standard deviations of the treatment and control group at endline were similar[18], the treatment

[18] Specifically, the difference in standard deviations at baseline was 0.6 for Section A and this difference was 0.9 at endline. It is expected that the standard deviation of Section B scores, had Section B been written at baseline, would have been greater in the treatment group. So because the distributions are so different, it is difficult to speculate

distribution of Section B results is much flatter at endline, corresponding to a higher standard deviation in the treatment group of 19.4 percent compared to a standard deviation of only 16.7 percent in the control group.

Figure 3: Distribution of mathematics scores – Grade 8 curriculum at endline



Source: Own data collected in 2013

The average effect of treatment is positive and highly significant with an increase in 5.8 percentage points for Section A and 3.9 percentage points on Section B in the treatment group when compared to the control group, holding all else constant.

---

how much of this difference in variance is due to baseline differences between control and treatment distributions in ability and how much is a result of the programme.

Drop-Out and Attendance Rates

Programme Drop-Out Rates

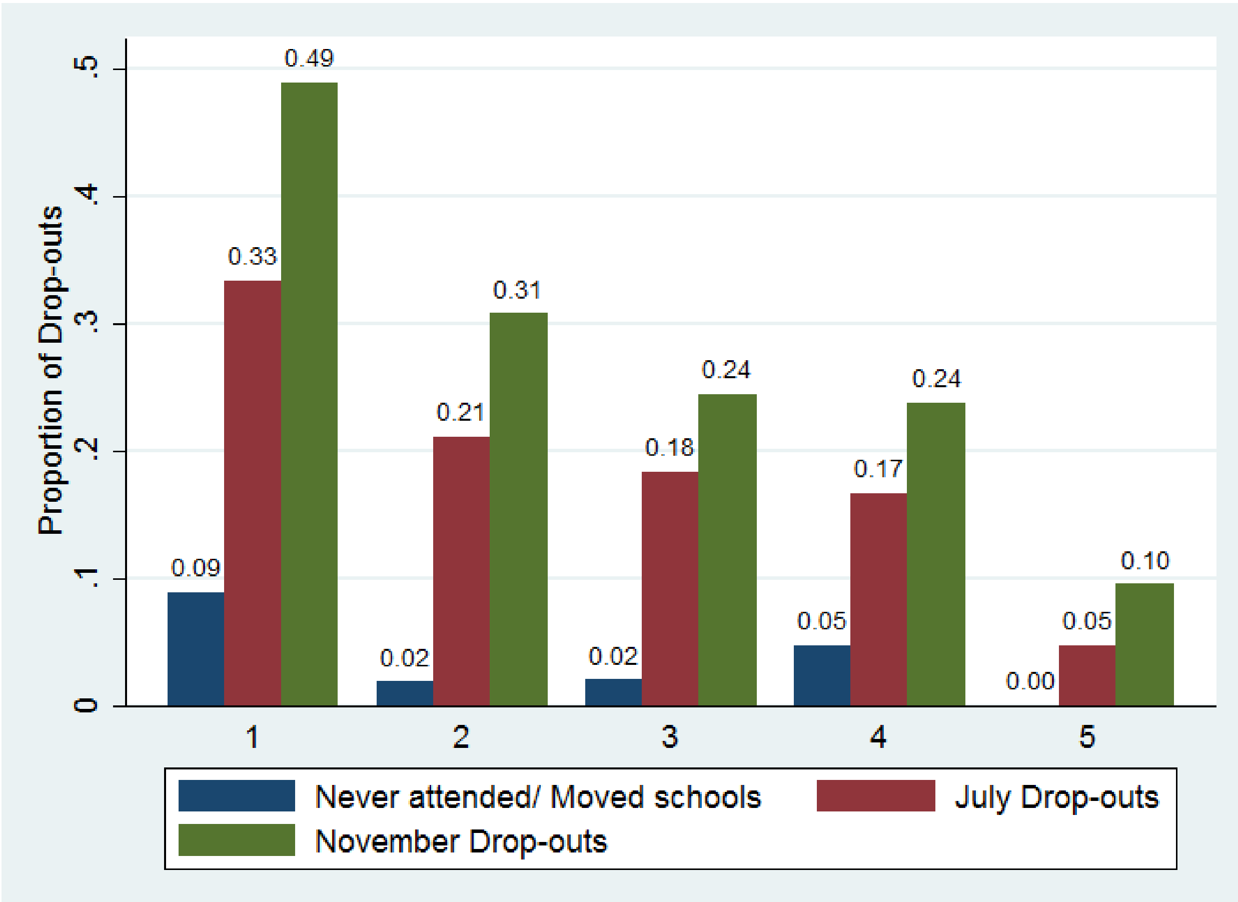
The success of the programme is closely linked to the real take-up rates among selected learners. Of the 236 learners that had originally been selected, there were 8 learners in total who did not attend a single class. Overall the drop-out rate by the end of the second term was moderate. A total of 191 (81 percent) learners were still actively attending classes. Over the next two terms, a further 23 learners dropped out of the programme. This equates to a drop-out rate of 19 percent over the first two terms and a total drop-out rate of 29 percent over the course of the year. Therefore, only 7 out of every 10 learners in the sample completed a full year of treatment.

> ***Result 5: Treatment learners that perform poorly on the baseline mathematics test are significantly more likely to drop out of the programme over the course of the year.***

In the results of an intention-to-treat analysis all learners selected into treatment are regarded as treated learners. However, in order for the treatment to have had any impact, learners need to have attended Numeric classes. Dropping out or ceasing to attend is a decision on the side of the learner that is decidedly non-random. Figure 4 shows the proportion of learners that dropped out of the programme by baseline mathematics quintiles[19]. It is evident that there is a strong negative correlation between achievement on the baseline mathematics test and drop-out rates. By the end of July, 33 percent of learners from the bottom quintile had stopped attending Numeric classes, whereas only 5 percent of learners from the fifth quintile had dropped out. In November, half of the learners in the bottom quintile were no longer attending classes compared to only 10 percent of top quintile learners.

[19] Quintile 1 comprises of learners from the lowest 20 percent of the baseline mathematics distribution, whilst quintile 5 consists of the top performers at baseline.

Figure 4: Proportion of learners that did not attend any classes, dropped out by the end of the second term and those that had dropped out at the end of the 4th term by baseline numeracy quintile



Source: Own data collected in 2013

Discussion

Introducing computer assisted learning into schools is one of the proposed means of raising the quality of instruction in developing countries, by simultaneously addressing any specific gaps in knowledge that a learner might have as a result of absenteeism and disruptions during teaching time. The Numeric programme is one such programme that used CAL in the form of Khan Academy to revise and teach basic numeracy.

Summary of Results

The endline mathematics results showed that the Numeric programme was successful at improving the basic numeracy (Section A) of learners selected into the treatment group by an average of 6.0 percentage points (0.32 s.d.) above the mean scores obtained by learners in the control group. Thus, the chief aim of increased numeracy was achieved. Furthermore, learners in treatment also outperformed learners in the control group by a mean of 4.1 percentage points (0.25 s.d) in Section B, which tested Grade 7 and 8 curriculum material. In this manner CAL was being used to increase the rate at which later learning could occur. This marked improvement as a result of selection into treatment suggests that learners' mathematics ability has objectively improved, whilst learner feedback shows that learners are aware of this improvement in their mathematical ability. Furthermore, learner feedback shows that learners on the programme are also enjoying mathematics more than learners in the control group. CAL has the added benefit in that gamefication and instant feedback enhance the learning experience which could have resulted in a more enjoyable experience (Schonfield, Eurich-Fulcer & Britt, 1994).

Lessons Learned

The Numeric programme takes place after school and is not linked to the South African school curriculum. This differentiates Numeric from most computer assisted learning interventions. Essentially, the programme offered mathematics classes as an extra-curricular activity as opposed to incorporating it into the learner's school day as a school subject.

Education interventions, particularly computer based interventions, by and large take the form of in-school programmes. These then act as a supplement for traditional teaching and would be held during school hours. On the other hand, an out-of-school or after-school programme, such as Numeric, takes place outside of school hours. In South Africa the proportion of learners who progress to secondary school is reasonably high[20]. However, fundamental skills and knowledge of basic numeracy are often lacking. It would be implausible to expect Grade 8 mathematics teachers to re-teach concepts covered in the Grade 3 or Grade 5 curriculum, although many of the learners in their classes may not have fully comprehended the concepts. Thus, the introduction of an after-school programme for older learners solves this problem. Since it takes place outside of school, the learners and coaches are not constrained by the school curriculum. Simultaneously, learners are not at risk of falling further behind in mathematics as they are still attending their regular school classes.

Limitations on the external validity of the results

There are limitations intrinsic to the design of the experiment and the nature of the programme. Two major concerns are the external validity of the results within South Africa and the potential change in the treatment effect as a result of the scaling up of the programme.

The external validity of the results is limited due to a number of factors. The Cape Metro Area is not representative of the country as a whole, and whilst it is well suited as a testing ground, it is not immediately obvious that similar effect sizes would be expected in other parts of the country, or even in more rural parts of the province. The schools selected for the pilot were wealthier than the average South African school, however, the results from this sample of 9 schools are encouraging. The difference between treatment and control at endline on Section A of both quintile 3 schools was 10 percentage points, which is 4 percentage points greater than the average treatment effect, whilst the difference in Section B scores between treatment and control were also higher than the average of 4.1 percentage points (see Table 17 and Table 5).

It is not clear whether treatment would be more or less effective in a rural as opposed to an urban setting. A number of factors may play a part in increasing or decreasing the effectiveness of the programme. Further research would need to be conducted in order to establish how effective the

20 The Department of Basic Education (2011b) in a Dropout and Learner Retention Strategy report states that 2.7 percent of students enrolled in Grade 7 in 2007 dropped out between 2007 and 2008. The same report claims that 5.1 percent of Grade 7 students were repeating Grade 7 in 2009. Noting that the original sample of Grade 7's would have included both repeating and non-repeating students, this puts the lower bound of students who progress to high school at 94.6 percent (under the assumption that the only learners that drop out of Grade 7 are repeating learners). Realistically, this number is unlikely to be much higher than about 3 or 4 percent. Alternatively, between 81.6 and 90.9 percent of learners that enroll in Grade 1 complete Grade 7 and start Grade 8.

Numeric programme would be in a poor, rural setting. Additionally, increasing the sample size to test the intervention in multiple different locations would produce valuable information.

The pilot was only run in schools which already had functional, but underutilised, computer laboratories. The cost and time implications of setting up the appropriate infrastructure are huge, and it is likely that the setup cost could significantly decrease the cost effectiveness of the programme in schools or communities where the computer facilities and internet connectivity do not yet exist. However, the setup of such a facility would most likely provide to school with benefits besides improved numeracy. Future studies would need to examine what the cost implications and feasibility of such an extension to the programme would be.

The Numeric programme in its current form is a small, privately run, non-profit organization. As a result of the relatively small size, the level of oversight, motivation and quality of staff, follow up with learners, teachers and parents and general management is of a high standard. Thus, there are some concerns with regards to scalability. An example of such a case is the Perry Preschool Project. This randomized trial was conducted in the 1960's and positive effects on learner achievement were still evident up until the age of 27 (Barnett, 2011). However, large scale public programmes utilising the same methods failed to achieve any gain in cognitive outcomes. One explanation is that the efficacy of the programme depends on the quality of implementation, which may be affected when scaling takes place, especially when the scaling up comes about as a result of a large-scale government intervention. The programme is currently in the form of an after-school programme that is in all probability complementing what teachers are teaching in the classroom. Additionally, learners needed to apply to participate as it was an optional after-school activity. There is no evidence to suggest similar improvements could be expected of an in-school programme targeting all learners which adopted a similar strategy to the Numeric after-school programme.

A last caveat in the interpretation of the main result is that the average impact on standardised mathematics scores was as a result of the Numeric programme as a whole. No attempt was made to determine which aspects of the programme were responsible for the improvement on mathematics scores. Further, evaluation would be required to determine which elements or combinations of features are most essential for improving the mathematics outcomes of learners.

Conclusion

There is an urgent need in South Africa to discover ways in which to improve the quality of education which is on offer for learners who are currently at school, as well as unemployed and unenrolled youths. The combination of an after-school programme using computer assisted learning proved a powerful way to significantly improve the mathematics outcomes of secondary school learners in their first year of secondary school. The after-school nature of the programme allowed the implementing organization to effectively utilise the infrastructure and resources of the schools, whilst simultaneously maintaining the freedom to operate outside of the nationally determined curriculum. The use of CAL, furthermore, allowed for individualized learning and focused remediation that significantly improved the basic numeracy of the learners on the programme, whilst simultaneously also improving their outcomes on core Grade 8 curriculum. Overall, the Numeric programme is a positive step towards addressing one of the problems facing learners in South African secondary schools, namely, a poor foundation in mathematics.

References

Angrist, J. & Lavy, V. 2002. New Evidence on Classroom Computers and Pupil Learning. *The Economic Journal*. 112: 735-765.

Banerjee, A. V., Cole, S., Duflo, E., & Linden, L. 2007. Remedying education: Evidence from two randomized experiments in India. *The Quarterly Journal of Economics*. 122(3): 1235-1264.

Banerjee, A. V. & Duflo, E. 2009. The Experimental Approach to Development Economics. *Annual Review of Economics.* 1:151–78

Barnett, W. S. 2011. Effectiveness of Early Educational Intervention. *Science*. 333: 975-978.

Barrara-Osorio, F. & Linden, L. L. 2009. The Use and Misuse of Computers in Education: Evidence from a Randomized Experiment in Colombia. World Bank Policy Research Working Paper 4836.

Barro, R. J. & Lee, J. J. 2001. International data on educational attainment: updates and implications. *Oxford Economic Papers*. 53(3): 541-563.

Barro, R. J. 2001. Human Capital and Growth. *The American Economic Review*. 91(2): 12-17.

Barrow, L., Markman, L. & Rouse, C. E. 2009. Technology's Edge: The Educational Benefits of Computer-Aided Instruction. *American Economic Journal: Economic Policy.* 1(1): 52-74.

Burde, D. & Linden, L. L. 2012. The Effect of Village Based Schools: Evidence from a Randomized Controlled Trial in Afghanistan. National Bureau of Economic Research Working Paper 18039.

Cristia, J. P., Ibarraran, P., Cueto, S., Santiago, A. & Severin, E. 2012. Technology and Child Development: Evidence from the One Laptop per Child Programme. Inter-American Development Bank Working Paper No. IDB-WP-304.

Cross, M., Mungadi, R., & Rouhani, S. 2002. From policy to practice: Curriculum reform in South African education. *Comparative Education*. 38(2):171-187.

Department Of Basic Education. 2004. White Paper on e-Education. Governmente Gazette, No. 26762 , 2 September 2004. Available: http://www.education.gov.za/DocumentsLibrary/Legislation/WhitePapers/tabid/191/Default.aspx [2014, 21 January].

Department Of Basic Education. 2010. The status of the language of learning and teaching (LOLT) in South African public schools: A quantitative overview. Available:

http://www.education.gov.za/LinkClick.aspx?fileticket=LlfRGMZxPRg%3D&tabid=422&mid=1261 [2014, 14 February].

Department Of Basic Education. 2011a. Curriculum News May 2011. Available: http://www.education.gov.za/LinkClick.aspx?fileticket=RlQ3WgihTOA%3D&tabid=348&mid=1018 [2014, 28 January].

Department Of Basic Education. 2011b. Report on Dropout and Learner Retention Strategy. Report to the Portfolio Committee on Education in June 2011, Available: http://www.education.gov.za/LinkClick.aspx?fileticket=jcSsY0rHcME%3D&tabid=358&mid=1261 [2014, 11 February].

Department Of Basic Education. 2013a. Annual Report 2012/2013. Pretoria: Department of Basic Education.

Department Of Basic Education. 2013b. Education Statistics in South Africa 2011. [Online data file]. Available: http://www.education.gov.za/EMIS/StatisticalPublications/tabid/462/Default.aspx [2013, 21 January].

Department Of Basic Education. 2013c. Masterlist data - March 2013 Ordinary Schools: Western Cape Province [Online data file]. Available: http://www.education.gov.za/EMIS/EMISDownloads/tabid/466/Default.aspx [2013, 23 October].

Fairlie, R. W. & Robinson, J. 2012. Experimental Evidence on the Effects of Home Computers on Academic Achievement among Schoolchildren. Unpublished.

Finn, A., Leibbrandt, M., & Woolard, I. 2009. Income and expenditure inequality: Analysis of the NIDS Wave 1 dataset. National Income Dynamics Study. NiDS Discussion Paper 5.

Galston, W. A. 2001. Political knowledge, political engagement, and civic education. *Annual Review of Political Science.* 4(1): 217-234.

Glass, G. V. 1977. Integrating Findings: The Meta-Analysis of Research. *Review of Research in Education*. 5: 351-379.

Glewwe, P, Kremer, M & Moulin, S. 2002. Textbooks and Test Scores: Evidence from a Prospective Evaluation in Kenya. Poverty Action Lab Paper No. 8.

Glewwe, P. 2002. Schools and Skills in Developing Countries: Education Policies and Socioeconomic Outcomes. *Journal of Economic Literature*. 40(2): 436-482.

Glewwe, P. & Kremer, M. 2002. Schools, Teachers, and Education Outcomes in Developing Countries. *Handbook of the Economics of Education*. 2: 945-1017

Gonzales, P., Williams, T., Jocelyn, L., Roey, S., Kastberg, D., & Brenwald, S. (2008). Highlights from TIMSS 2007: Mathematics and Science Achievement of US Fourth-and Eighth-Grade Students in an International Context. (Research report; NCES 2009-001). Washington,DC, U.S: US National Center for Education Statistics.

Hembree, R. 1990. The Nature, Effects, and Relief of Mathematics Anxiety. *Journal for Research in Mathematics Education*. 21(1): 33-46.

John, V. 2014. Matric pass rate highest yet at 78.2%. Mail and Guardian Online. Available: http://mg.co.za/article/2014-01-06-matric-pass-rate-highest-yet-at-782 [Accessed on 21 January 2014].

Khan Academy. 2014. Khan Academy FAQ. Available: http://khanacademy.desk.com/ [2014, 21 January].

Kremer, M. & Holla, A. 2009. Improving Education in the Developing World: What Have We Learned from Randomized Evaluations?. *Annual Review of Economics.* 1(1): 513-542.

Kremer, M., Brannen, C., & Glennerster, R. 2013. The challenge of education and learning in the developing world. *Science*, 340(6130): 297-300.

Linden, L. L. 2008. Complement or Substitute? The Effect of Technology on Student Achievement in India. Available: http://www.poverty-action.org/node/2241 [2013, 18 August].

Louw, J., Muller, J. & Tredoux, C. 2008. Time-on-task, technology and mathematics achievement. *Evaluation and Program Planning*. 31: 41-50.

Malamud, O. & Pop-Eleches, C. 2010. Home Computer Use and the Development of Human Capital, National Bureau of Economic Research Working Paper 15814.

Mankiw, N. G., Romer, D. & Weil, D. N. 1992. A Contribution to the Empirics of Economic Growth. *The Quarterly Journal of Economics.* 107(2): 407-437.

Mckenzie, D. 2012. Beyond baseline and follow-up: The case for more T in experiments. *Journal of Development Economic*. 99(2): 210-221.

Miguel, E. & Kremer, M. 2004. Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities. *Econometrica*. 72: 159-217.

Mybroadband (2014). Broadband price per Mbps shocker for South Africa (11 January 2014), Available at: http://mybroadband.co.za/news/broadband/93895-broadband-price-per-mbps-shocker-for-south-africa.html [Accessed 14 February 2014].

Ramatapana, K. & Makonye, J. P. 2012. From too much freedom to too much restriction: The case of teacher autonomy from National Curriculum Statement (NCS) to Curriculum and Assessment Statement (CAPS). *Africa Education Review*. 9(1): S7-S25

Raven, J., Raven, J. C. & Court, J. H. 2003. Raven Manual: Section 3. San Antonio: Pearson

Roschelle, J. M., Pea, R. D., Hoadley, C. M., Gordin, D. N. & Means, B. M. 2000. Changing How and What Children Learn in School with Computer-Based Technologies. The Future of Children, Children and Computer Technology. 10(2): 76-101.

Schonfield, J. W., Eurich-Fulcer, R. & Britt, C. L. 1994. Teachers, Computer Tutors, and Teaching: The Artificially Intelligent Tutor as an Agent for Classroom Change. *American Educational Research Journal*. 31(3): 579-607.

Spaull, N. 2011. Primary School Performance in Botswana, Mozambique, Namibia, and South Africa. SAQMEQ Working Paper 8, September 2011.

Spaull, N. 2013. South Africa's Education Crisis: The quality of education in South Africa 1994-2011. (CDE report October 2013). Johannesburg, South Africa: Center for Development and Enterprise

Speakman, R. & Welch, F. 2006. Chapter 13: Using Wages to Infer School Quality. In Kenneth J. Arrow and Michael D. Intriligator (eds). Handbook of the Economics of Education. 2: 818-864.

Statistics South Africa. 2010a. Census At School Results 2009, Available http://www.statssa.gov.za/censusatschool/docs/Census_At_School_2009_Report.pdf [2014, 21 January].

Statistics South Africa. 2010b. Republic of South Africa Millenium Development Goals Country Report 2010. Available: http://www.info.gov.za/view/DownloadFileAction?id=132022 [2013, 17 August].

Taylor, S. 2011. Uncovering indicators of effective school management in South Africa using the National School Effectiveness Study. Ph.D. Thesis. University of Stellenbosch.

Van der Berg, S. 2006. The targeting of public spending on school education, 1995 and 2000. *Perspectives in Education: Education Finance*. 24(2): 49.

Van der Berg, S. 2007. Apartheid's enduring legacy: Inequalities in education. *Journal of African Economies*. 16(5): 849-880.

Western Cape Education Department. 2009. WCED celebrates 1000 Khanya schools, Media Release. Available: http://wced.pgwc.gov.za/comms/press/2009/41_khanya.html [2014, 21 January].

Western Cape Education Department. 2013a. Background to the national quintile system. 14 October 2013. Available: http://wced.pgwc.gov.za/comms/press/2013/74_14oct.html [2014, 11 February].

Western Cape Education Department. 2013b. Education management information systems: Find-a-school. Available: http://wcedemis.pgwc.gov.za/wced/findaschool.html [2013, 23 October].

White, I. R. & Thompson, S. G. 2005. Adjusting for partially missing baseline measurements in randomized trials. *Statistics in Medicine*. 24(7): 993-1007.

World Bank. 2014. World Development Indicators [Online data file]. Available: http://data.worldbank.org/country/south-africa [2014, 21 January].