
Starting Behind and Staying Behind in South Africa
The case of insurmountable learning deficits in mathematics

NICHOLAS SPAULL AND JANELI KOTZE

Stellenbosch Economic Working Papers: 27/14

KEYWORDS: MATHEMATICS, LEARNING TRAJECTORIES, SOUTH Africa,
HIERARCHICAL LEARNING, SACMEQ, TIMSS
JEL: I20, I21, I28

NICHOLAS SPAULL
DEPARTMENT OF ECONOMICS
UNIVERSITY OF STELLENBOSCH
PRIVATE BAG X1, 7602
MATIELAND, SOUTH AFRICA
E-MAIL: NICHOLASSPAULL@GMAIL.COM

JANELI KOTZE
DEPARTMENT OF ECONOMICS
UNIVERSITY OF STELLENBOSCH
PRIVATE BAG X1, 7602
MATIELAND, SOUTH AFRICA
E-MAIL: JANELI.VILJOEN@GMAIL.COM



UNIVERSITEIT
STELLENBOSCH
UNIVERSITY



A WORKING PAPER OF THE DEPARTMENT OF ECONOMICS AND THE
BUREAU FOR ECONOMIC RESEARCH AT THE UNIVERSITY OF STELLENBOSCH

Starting Behind and Staying Behind in South Africa

The case of insurmountable learning deficits in mathematics

Nicholas Spaul¹ & Janeli Kotze²

Abstract

This study quantifies a year's worth of mathematics learning in South Africa (0.3 standard deviations) and uses this measure to develop empirically-calibrated learning trajectories. Two main findings are, (1) only the top 16% of South African Grade 3 children are performing at an appropriate Grade 3 level. (2) The learning gap between the poorest 60% of students and the wealthiest 20% of students is approximately three Grade-levels in Grade 3, growing to four Grade-levels by Grade 9. The paper concludes by arguing that the later in life we attempt to repair early learning deficits in mathematics, the costlier the remediation becomes.

Keywords: Mathematics; Learning Trajectories; South Africa; Hierarchical Learning; SACMEQ; TIMSS

Highlights:

- Only 16% of Grade 3 students in South Africa are performing at a Grade 3 level in mathematics.
- The poorest 60% of students are three Grade-levels behind the wealthiest 20% of students in Grade 3.
- The gap between the poorest 60% and wealthiest 20% of students grows to four Grade-levels by Grade 9.

JEL Codes: I20, I21, I28

Acknowledgements: We would like to acknowledge the South African Human Rights Commission (SAHRC) who provided the funding for this research.

¹ Corresponding Author: Nicholas Spaul, NicholasSpaul@gmail.com

² Both authors are at the University of Stellenbosch, South Africa.

The authors would like to thank Professor Servaas van der Berg and the Research on Socioeconomic Policy (ReSEP) team at Stellenbosch for their feedback and suggestions. We would also like to thank two anonymous reviewers for their helpful suggestions and advice.

1. Introduction

Few would argue that the state of mathematics education in South Africa is something other than dire. This belief is widespread among academic researchers and those in civil society, and is also strongly supported by a host of local and international assessments of mathematical achievement extending back to at least 1995 (Howie & Hughes, 1998; Reddy, 2006; Fleisch, 2008; Spaul, 2013; Taylor et al., 2013). Many of these studies, and particularly those that focus on mathematics, have identified that students acquire learning deficits early on in their schooling careers and that these backlogs are the root cause of underperformance in later years. They argue that any attempts to raise students' mathematical proficiency must first address these deficits if they are to be successful (Taylor et al., 2003). The present study adds further evidence to this body of work by using nationally representative data to provide some indication of the true size and scope of these learning deficits.

Both Prichett & Beatty (2012) and Banerjee & Duflo (2011) have identified that students in developing countries have large learning deficits. They show that even children with relatively high levels of educational attainment often have very few cognitive skills to show for all their years of schooling. They theorise that this is the result of weaker students falling progressively further and further behind the curriculum to the extent that they eventually fall so far behind that no learning takes place whatsoever. Muralidharan and Zieleniak (2013) found support for this hypothesis with data from the Andhra Pradesh Randomized Evaluation Studies from India, by means of tracking the learning of a group of students over a five year period. Their results show that only 60% of students reach a Grade 1 level after five years of formal full-time schooling, and furthermore, that the learning trajectories of the weakest performers flatten off completely in the later Grades. This provides empirical support to Lewin's (2007, p.10) notion of 'silent exclusion' where students are enrolled and attending school but learning little.

In South Africa, research in this area has generally focussed on in-depth localized studies of student workbooks and classroom observation (Ensor *et al.*, 2009). For some examples, Carnoy *et al.* (2012) observe mathematics learning in Grade 6 classrooms from 60 schools in one South African province (North West) and compare these classrooms to 60 schools in neighbouring Botswana. On a smaller scale, Venkat & Naidoo (2012) focus on 10 primary schools in Gauteng and analyse coherence for conceptual learning in a Grade 2 numeracy lesson. Similarly Schollar (2008) conducted interviews and classroom observations as well as analysed a large sample of learner scripts to determine the development (or lack thereof) of mathematical concepts through the Grades.

Where the present research differs from these earlier studies is that it focuses on quantifying national learning deficits in general, rather than in specific learning areas. While the latter are essential for understanding what the problems are and how to fix them, analyses at the national level are also needed if we are to understand the extent and distribution of the problem, both of which are imperative for policy-making purposes. This is only possible by analysing multiple nationally-representative surveys of student achievement, which is the focus of the present study. The two core research questions that animate this study are as follows:

- 1) How large are learning deficits in South Africa and how are they distributed in the student population?
- 2) Do learning deficits grow, shrink or remain unchanged as students progress to higher Grades?

To answer these questions we analyse four nationally representative datasets of mathematics achievement, namely: (1) the Systemic Evaluation 2007 (Grade 3), (2) the National School Effectiveness Study 2007/8/9 (Grade 3, 4 and 5), (3) the Southern and Eastern African Consortium for Monitoring Educational Quality (SACMEQ) 2007 (Grade 6), (4) and the Trends in International Mathematics and Science Study (TIMSS) 2011 (Grade 9).

2. Background

Independent studies in economics, neuroscience and developmental psychology all confirm that the mastery of the skills which are essential for economic success and personal development largely follow hierarchical rules (Knudsen et al., 2006). The later acquisition of these skills builds on the foundations laid down in earlier years. That is, earlier mastery of certain cognitive, social and emotional capabilities help foster more efficient learning at later ages. Conversely, the lack of certain capabilities creates a low ceiling beyond which progress is improbable. Developing a theory of learning that incorporates these insights, Robert Gagné (1962) proposed the notion of 'learning hierarchies' as a set of ordered intellectual skills which are hierarchically inter-related. He posited that a final capability can be broken into subordinate skills in such a manner that lower-level capabilities generate a substantial amount of positive transfer to the learning of higher order capabilities that have not yet been acquired (see also Scandura & Wells, 1967). These theories have considerable empirical support, with numerous studies finding early numeracy skills to be good predictors of later mathematics performance (Aubrey & Godfrey, 2003; Aubrey et al., 2006; Aunio & Niemivirta, 2010). Counting skills, in particular, have been shown to estimate basic arithmetic skills in the early Grades of primary school relatively accurately (Aunola et al., 2004; Jordan et al., 2007; Desoete et al., 2009).

An epistemological analysis of mathematics reveals a latent hierarchy of knowledge and intellectual skill - what Posner and Strike (1976) refer to as content structure; "Content structure refers to the content elements and the ordering relationships that exist between them ... Most questions about content structure can be reduced to questions concerning what content comes before what other content and the rationale for that order" (Posner & Strike, 1976: p. 666, cited in Reeves & McAuliffe, 2012, p. 11). Consequently the acquisition of higher order knowledge and intellectual skills requires first the mastery of subordinate skills and a clear understanding of foundational mathematical knowledge. This implicit knowledge structure is made explicit in the sequencing and structuring of curricula where simple antecedents precede more complex concepts and ways of thinking. Kilpatrick, et al. (2001) encapsulates this concept of mathematical proficiency as five interwoven and interdependent strands. They explicitly state that mathematical proficiency cannot be attained by only focussing on one strand, but for students to progress in mathematical proficiency all strands need to be developed. Although this is true of many - if not most - subjects, mathematics is perhaps the best example of such a subject due to the strong vertical demarcation and integration of concepts. For example, without an understanding of the concepts of number and equipartitioning a student will not be able to understand or manipulate fractions which are necessary for fraction equivalence and comparison.

The extant research on mathematics learning in South Africa strongly supports this conclusion with numerous researchers highlighting the inadequate acquisition of basic skills and the consequent negative effects on further learning. Taylor & Vinjevoold (1999) summarise the findings from 54 studies³ commissioned by the President's Education Initiative and conclude that:

"At all levels investigated by [The President's Education Initiative], the conceptual knowledge of students is well below that expected at the respective Grades. Furthermore, because students are infrequently required to engage with tasks at any but the most elementary cognitive level, the development of higher order skills is stunted" (Taylor & Vinjevoold, 1999, p. 231).

This lack of engagement with higher order content is the prime focus of Reeves and Muller's (2005) analysis of Opportunity-to-Learn (OTL) and mathematics achievement in South Africa, where OTL is the curriculum actually made available to learners in the classroom. Taylor et al. (2003, p. 129) in their book *Getting Schools Working* summarise succinctly the debilitating effects of cumulative learning deficits:

"At the end of the Foundation Phase [Grades 1-3], learners have only a rudimentary grasp of the principles of reading and writing ... it is very hard for learners to make up this cumulative deficit in later years ... particularly in those subjects that ... [have] vertical demarcation requirements (especially mathematics and science), the sequence, pacing, progression and coverage requirements of the high school curriculum make it virtually impossible for learners who have been disadvantaged by

³ The full list of research papers can be found at <http://www.jet.org.za/publications/pei-research>

their early schooling to 'catch-up' later sufficiently to do themselves justice at the high school exit level.'

And lastly, Schollar (2008) summarises the findings of the Primary Mathematics Research Project which looked at over 7000 learners from 154 schools in South Africa and concludes as follows:

"Phase I concluded that the fundamental cause of poor learner performance across our education system was a failure to extend the ability of learners from counting to true calculating in their primary schooling. All more complex mathematics depends, in the first instance, on an instinctive understanding of place value within the base-10 number system, combined with an ability to readily perform basic calculations and see numeric relationships ... Learners are routinely promoted from one Grade to the next without having mastered the content and foundational competences of preceding Grades, resulting in a large cognitive backlog that progressively inhibits the acquisition of more complex competencies. The consequence is that every class has become, in effect, a 'multi-Grade' class in which there is a very large range of learner abilities and this makes it very difficult, or even impossible, to consistently teach to the required assessment standards for any particular Grade. Mathematics, however, is an hierarchical subject in which the development of increasingly complex cognitive abilities at each succeeding level is dependent on the progressive and cumulative mastery of its conceptual frameworks, starting with the absolutely fundamental basics of place value (the base-10 number system) and the four operations (calculation)" (Schollar, 2008, p. 1).

However, few of these studies use nationally representative samples in their analysis of student achievement, and none when looking specifically at learning deficits. This is not to say that there have not been a number of reports that have looked at the nationally representative datasets of educational achievement in South Africa⁴. However, these reports do not focus on learning deficits but rather the levels and trends of performance in the country. In this sense there is a bifurcation in the literature where small-scale studies focus on learning deficits without being able to make population-wide claims, while large-scale studies which can make population-wide claims do not look specifically at learning deficits.

It is important to mention that the term 'learning deficit' is used throughout this paper to foreground the absence of learning and remedial opportunities, not an inherent learning deficit or disability of the child. We take the term learning deficit to mean the difference between the actual performance of a student and some benchmark which is used as a reference category. In no way should this term be interpreted as referring to an individual child's ability or lack thereof. With the exception of those with neurological impairments, we take the view that all children – rich and poor alike – have equal innate potential to acquire knowledge, skills and values. Too often in the South African educational discourse the performance of poorer students is wrongly seen as some reflection of the underperforming child rather than some reflection of the underperforming system.

⁴ For some examples see Fiske & Ladd, 2004; Reddy, 2006; Fleisch, 2008; Taylor & Yu, 2009; Van der Berg, et al., 2011; Moloi & Chetty, 2011; and Spaull, 2013.

3. Data

In order to construct the learning trajectories of South African children, it is necessary to have objective measures of achievement at multiple points in the education system. For the present analysis we use three data sets which cover five Grades from Grade 3 to Grade 9. The data are drawn from the National School Effectiveness Study (NSES) for Grades 3, 4 and 5; from the Southern and Eastern African Consortium for Monitoring Educational Quality (SACMEQ) for Grade 6; and from the Trends in International Mathematics and Science Study (TIMSS) for Grade 9. Given that we also discuss the Systemic Evaluation of 2007 (Grade 3) we also provide background information for that data set.

3.1 Systemic Evaluation – Grade 3 (2007)

The 2007 Systemic Evaluation tested a nationally representative sample of Grade 3 students in numeracy and literacy. A random sample of about 54 000 Grade 3 students from 2 340 primary schools participated in the study (DoE, 2008). These students were assessed through standardised literacy and numeracy tests which measured their levels of achievement in terms of the Grade appropriate curriculum. To achieve this measure, the test comprised of Grade 1 to Grade 4 level questions, with the vast majority being set at the Grade 3 level. The tests were administered in all 11 official South African languages according to the Language of Learning and Teaching (LoLT) specified by the school.

3.1 National School Effectiveness Study (NSES) – Grade 3 (2007), Grade 4 (2008), Grade 5 (2009)

The NSES study is the first nationally representative⁵ panel data set which focusses specifically on schooling and educational outcomes (see Taylor et al. (2013) for a full discussion). The panel followed one cohort of students and tested them in Grade 3 (2007), Grade 4 (2008) and Grade 5 (2009). Approximately 15 000 students from 266 schools were tested each year with 8 383 students matched consistently across the three years and 24 000 tested in total across the three years. In this paper the 8 383 students who were observed in all three years are referred to as the panel sample, while the full 24 000 students are referred to as the full sample. The students wrote the exact same literacy and numeracy tests in each consecutive year, thereby producing vertically scaled, comparable results over time. Both the literacy and numeracy test paper were exact replicas of the Systemic Evaluation (2007) test papers, with the exception that the NSES was administered only in English. The questions included in the numeracy test ranged from Grade 1 to Grade 4 level, specified according to the National Curriculum Statement (NCS).⁶ Additional

⁵ Gauteng was not included in the NSES sample due to the fact that other testing was being conducted in that province at the same time.

⁶ The National Curriculum Statement is the curriculum which was taught in schools from 2002 until 2009 (Department of Education, 2002).

information with regards to student background, teacher characteristics and school principal characteristics were also collected over those years.⁷

3.2 Southern and Eastern African Consortium for Monitoring Educational Quality (SACMEQ) – Grade 6 (2007)

SACMEQ is a consortium of education ministries, policy-makers and researchers who, in conjunction with UNESCO's International Institute for Educational Planning (IIEP), aims to improve the research capacity and technical skills of educational planners in participating countries in Africa (Murimba, 2005; Moloji & Strauss, 2005: 12). These surveys collect extensive background information on the schooling and home environments of Grade 6 students, and in addition, test students and teachers in both numeracy and literacy (see Ross, et al., 2005 and Hungi, et al., 2010). Currently there are 15 participating countries including South Africa. The data set used for the present analysis is SACMEQ III (2007) South Africa, which tested 9 071 Grade 6 students from 392 schools, forming a large nationally representative sample (Moloji & Chetty, 2011). The SACMEQ tests are constructed to be aligned with the curricula covered in the participating countries, and are therefore closer to the Grade appropriate level than large scale international tests.

3.3 The Trends in International Mathematics and Science Study (TIMSS) – Grade 9 (2011)

TIMSS is a cross-national study which tests the mathematics and science knowledge of Grade 8 students in over 60 countries in such a way that they are comparable across countries and over time (Mullis et al., 2012). In the 2002 TIMSS, South Africa tested Grade 9 students in addition to Grade 8 students, since earlier rounds of TIMSS indicated that the international Grade 8 test was too difficult for South African students, and consequently too many students were performing at guessing level on the multiple choice questions (i.e. no better than random). This decreases the reliability and accuracy of the tests (Foy et al., 2010) and thus in 2011, only Grade 9 South African students wrote the TIMSS Grade 8 test. In TIMSS 2011 South Africa tested a nationally representative sample of 11 969 Grade 9 students from 285 schools in both mathematics and science (Reddy, et al., 2012).

⁷ Teacher questionnaires, however, were only administered in the years 2008 and 2009.

4. Learning Deficits

4.1. The South African case

Given the cumulative nature of learning deficits, it seems logical to determine when these learning deficits arise, as well as their size and distribution in the student population. In an ideal world one would have longitudinal data on the social, emotional and cognitive skills of children before they enter school and then follow these same children as they progress through school, assessing them at each Grade. Such data would allow for the disaggregation of learning deficits and indicate which portion of the deficit is from a child's home background and which portion is from the child's schooling experiences (see, for example, the Early Childhood Longitudinal Study by the Institute of Educational Science in the USA). Unfortunately this data does not yet exist in South Africa. Although the recent Annual National Assessments (ANAs) tested all South African Grade 1-6 students in 2011, 2012, 2013 and 2014, these tests were not psychometrically calibrated to be comparable across time or between Grades making them unusable for comparison purposes (see Spaul, 2013 for a full discussion). Although there are no reliable nationally-representative data sets for Grades 1 or 2, there are two studies which tested nationally representative samples of Grade 3 students in 2007. The first of these studies was the 2007 Systemic Evaluation (SE) conducted in September 2007 and the second study, the National School Effectiveness Study (NSES) tested a sub-sample of students from the Systemic Evaluation one month later in October 2007. The only difference between the Systemic and NSES Grade 3 tests was that the Systemic Evaluation (September) was conducted in the language of learning and teaching (LOLT) of the school – i.e. all 11 official languages - whereas the NSES test (October) was conducted only in English.

The school language policy in South Africa is currently implemented in such a way that the language of learning and teaching (LOLT) for the vast majority of students is their home-language for Grades 1, 2 and 3 and that from Grade 4 there is a LOLT switch to English for the remaining school years (Taylor & Coetzee, 2013).⁸ Given that the Grade 3 Systemic Evaluation of 2007 was conducted in the language of learning and teaching of the school, this should provide an accurate reflection of the state of mathematics learning at the Grade 3 level and minimize any confounding factors arising from not being proficient in the English language. The Grade 3 Systemic Evaluation mathematics test consisted of 53 questions which varied according to the nature of the mathematical tasks, the difficulty level of the items, whether the item was in verbal or symbolic form, and whether the item was multiple choice or free response (Taylor & Taylor, 2013). Furthermore, the question items were also classified by learning area and Grade-level in accordance with the prevailing curriculum, the National Curriculum Statement (NCS). Of the 53

⁸ English and Afrikaans students learn in their home-language from Grade 1 to Grade 12 and do not switch in Grade 4. For a full discussion of the school language dynamics in South Africa see Taylor & Coetzee (2013).

questions in the test, three were set at a Grade 1 level, 14 at a Grade 2 level, 30 at a Grade 3 level, and six at a Grade 4 level. Using this information we calculate the average numeracy score for each child using only the subset of 30 Grade-3 level questions. The reason for this sub-classification by Grade level is that this provides us with information on the expected ability level at each Grade, which subsequently allows us to calculate the proportion of students that are performing at the Grade-appropriate level in Grade 3. Following Muralidharan & Zieleniak (2013) we classify students as performing at the Grade-appropriate level if they obtain a mean score of 50% or higher on the full set of Grade 3 level questions. Since the questions were marked as correct or incorrect, the mean score indicates the percentage of questions a student managed to answer correctly. A mean score of 50% therefore suggests that a student has a 50% chance of either answering a Grade appropriate question correct or incorrect, which in turn indicates that the student is performing at a Grade appropriate level.

Figure 1 below shows the distribution of mean Grade 3 performance on Grade 3 level items disaggregated by quintile of student socioeconomic status into the wealthiest 20% of students (Quintile 5) and the poorest 80% of students (Quintile 1-4). All students achieving a mean score of 50% or higher can be said to be performing at the Grade-appropriate level. The graph reveals the dire situation in South Africa where the vast majority (88%) of Quintile 1 – 4 students in Grade 3 are not performing at the Grade-appropriate level. Looking at the distribution of Quintile 1-4 students, it becomes clear that these students are substantially behind the benchmark (50%). The majority of Quintile 1-4 students are concentrated around the 20% performance mark, a full one and a half standard deviations below the 50% threshold. Although Quintile 5 students perform much better than their poorer counterparts, only slightly more than half (51%) are performing at the Grade-appropriate level (see Table 1 below).

Figure 1: Kernel Density of Mean Grade 3 Performance on Grade 3 Level Items for Quintile 1-4 (poorest 80% of students) and Quintile 5 (wealthiest 20% of students) (Systemic Evaluation, 2007)

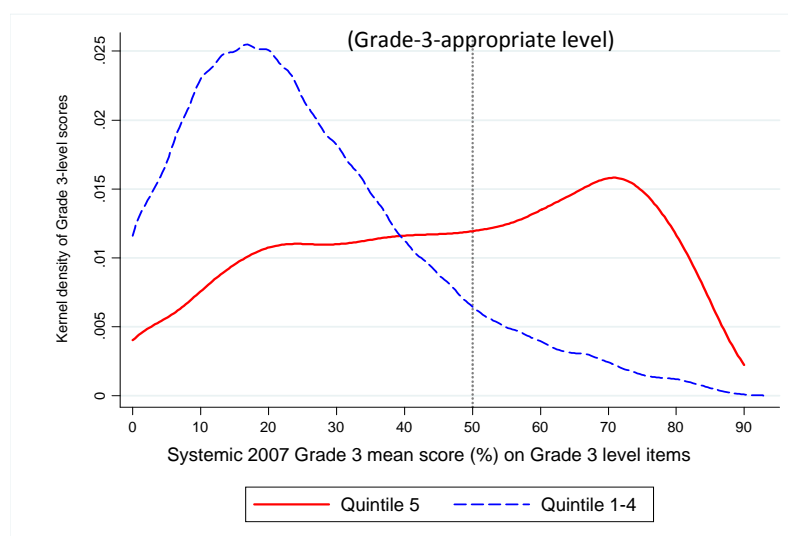


Table 1: Proportion of Grade 3 Students Performing at the Grade 3 Level by Province and Student Socioeconomic Quintile (Systemic Evaluation 2007)

Province	Proportion of Grade 3 students performing at the appropriate Grade 3 level*	Quintile	Proportion of Grade 3 students performing at the appropriate Grade 3 level*
Eastern cape	17%	Quintile 1	10%
Free State	25%	Quintile 2	10%
Gauteng	26%	Quintile 3	12%
KwaZulu-Natal	13%	Quintile 4	29%
Limpopo	6%	Quintile 1-4	11%
Mpumalanga	11%	Quintile 5	51%
North West Province	10%		
Northern Cape	17%		
Western Cape	32%		
South Africa	16%		

* Students are classified as performing at the Grade-appropriate level if they obtain a mean score of 50% or higher on the full set of Grade 3 level questions. Quintile 1 is the poorest 20% of students and Quintile 5 is the wealthiest 20% of students.

If one looks at the country as a whole, less than one in five (16%) Grade 3 students are performing at the Grade 3 level. That is to say that only the top 16% of Grade 3 students are performing at the Grade 3 level. Importantly, these Systemic assessments were conducted in the language of learning and teaching (LOLT) of the school in Grade 3, i.e. before any switch to English in Grade 4.

It is indisputable that by Grade 3 there already exist large learning deficits such that the vast majority of South African students (eight year olds) are well behind the curriculum. However, the origin of these learning deficits is less clear. Without longitudinal data on student achievement which covers the period before and during primary school, one cannot determine the source of these deficits, i.e. are they primarily attributable to having a disadvantaged home background, weak early childhood development or weak instruction in Grades 1, 2 and 3? Although we cannot answer this question with the data available in South Africa, we can answer another important and related question; whether learning deficits grow, shrink or remain constant as students' progress through the schooling system. To answer this question one needs to look at surveys of student performance at multiple points in the education system.

4.2. Learning Deficits in Grades 3, 4 and 5

One of the major nationally-representative datasets of student achievement in South Africa - and the only educational panel dataset in the country - is the National School Effectiveness Study (NSES) covering Grades 3, 4 and 5. All NSES tests were written in English only. Given the complex language dynamics in South Africa, with most students switching language in Grade 4, we chose to sub-classify the items in the mathematics test into "high-language" items and "no-

language” items⁹. An item was said to be a “high language” item if it was practically impossible to solve the problem without an understanding of the English language, whereas items were classified as “no language” items if they required no proficiency in the English language to solve them (i.e. they were entirely in number/symbol format). Of the 53 questions in the test 12 items¹⁰ had high language content and 15 items¹¹ had no language¹² content.

By focussing on the ‘no-language’ items and observing how students perform on these items as they progress from Grade 3 to Grade 5 it is possible to isolate the effect of increased mathematical proficiency from any confounding language factors arising from not being proficient in English. If we use the 50%-on-Grade-3-level-items threshold as a measure of the proportion of students operating at a Grade 3 level (as in Figure 1 above), and now also impose the “no language” restriction, we are left with nine items. In Panel 1 of Figure 2 below, only 8% of Grade 3 students from Quintile 1-4 were performing at the Grade 3 level according to these nine items. By contrast, 35% of Quintile 5 students were performing at the Grade-appropriate level. The second panel of Figure 2 shows that by Grade 5 this figure has increased substantially to 26% for Quintiles 1-4 and 55% for Quintile 5 students. It is disconcerting to note that only one in four (26%) Grade *five* students from Quintile 1-4 were operating at a Grade *three* level in 2009, at least according to these nine items, and furthermore that 45% of the wealthiest students (Quintile 5) are still not operating at a Grade 3 level by the end of Grade 5.

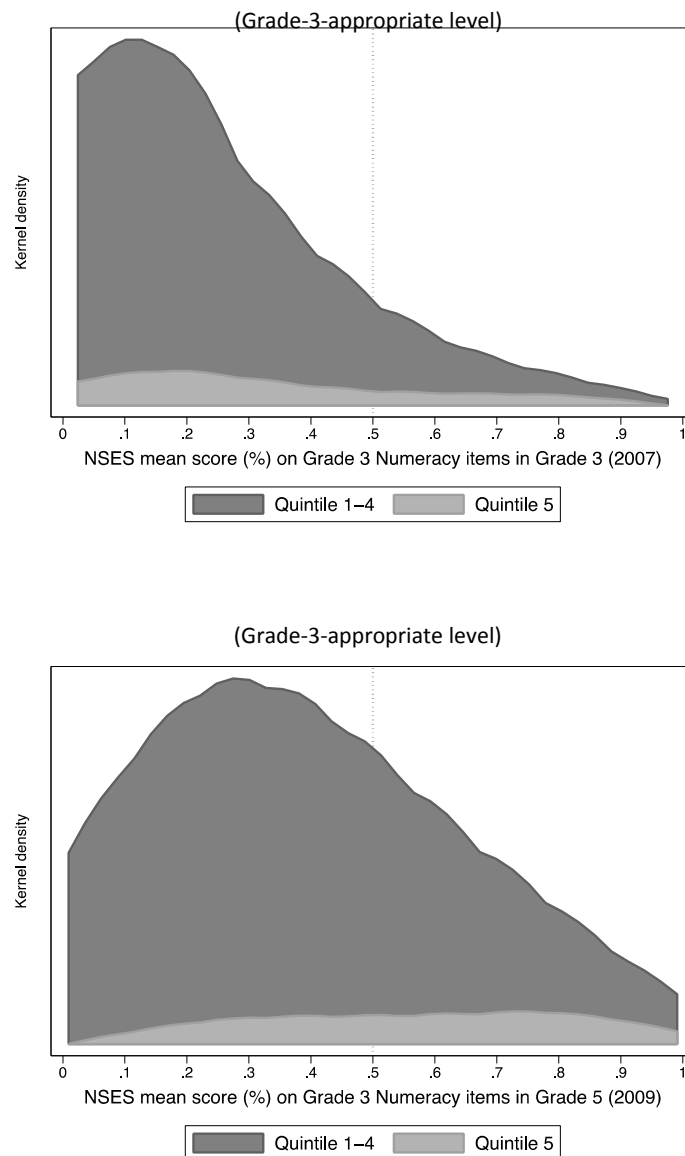
⁹ We recognise mathematic proficiency also requires the skill of reading, albeit in symbolic terms. The intention of this distinction is to control for any confounding factors flowing from students not being proficient in the language the test was administered in.

¹⁰ These 12 items were questions 10, 19, 22, 27, 29, 30, 33, 38, 43, 51, 52 and 53.

¹¹ These 15 items were questions 20, 21, 24, 25, 26, 28, 31, 32, 35, 36, 37, 39, 40, 41 and 42.

¹² We use the term language here in a limited sense to refer to English words and sentences made of words. However, we do acknowledge that in the broader sense language does also include reading and writing the symbolic forms that exist in mathematics.

Figure 2: NSES Grade 3 (panel 1) and Grade 5 (panel 2) performance on no-language items by quintile of student socioeconomic status (weighted and overlaid – full sample)



The above graphs clearly show that the majority of South African children are underperforming relative to the Grade-appropriate curriculum. However, such aggregated measures make it difficult to appreciate just how low the levels of performance really are, and how little learning occurs over the three years from Grades 3 to 5. To provide an alternative measure of performance, we provide two examples of no-language items in NSES and show when students answer the question correctly – i.e. in Grade 3, Grade 4, Grade 5 or not by the end of Grade 5. Given that one needs to follow the same students from Grade 3 to 5 we limit the sample here to the panel sample of NSES students (8383 students). Figure 3 below shows a simple question testing two and three digit addition with no carrying. This is within the Grade 3 curriculum which states that students should be able to “perform calculations using the appropriate symbols to solve problems involving addition of whole numbers with at least three digits.” Although this is a Grade 3 level item and contains no

language content, only 20% of Quintile 1-4 students could answer this correctly in Grade 3, with the proportion in Quintile 5 being twice as high (42%) but still low. While there is evidently some learning taking place in Grade 4 and 5, more than 40% of Quintile 1-4 children still could not answer this Grade 3 level problem at the end of Grade 5. In Quintile 5 this figure was only 22%.

Figure 3: National School Effectiveness Study – Proportion of students answering the item correctly by Grade and quintile of student socioeconomic status (NSES Question 21)

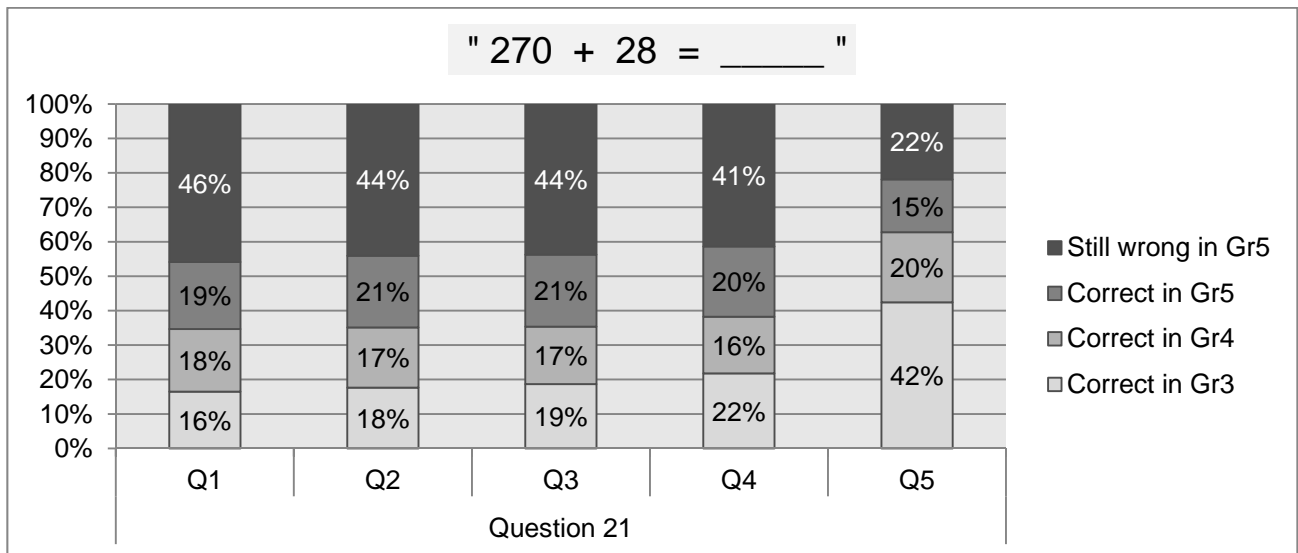
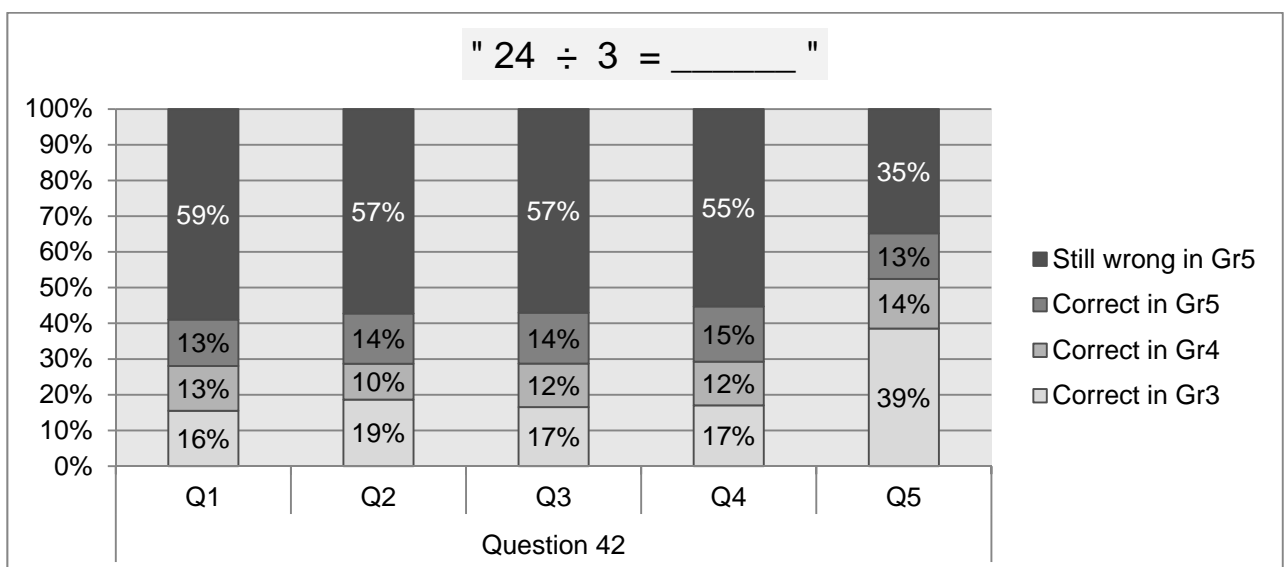


Figure 4 below shows a similar situation where the vast majority of Grade 3 children cannot answer this Grade 3 level problem. While some children learn the skill in Grade 4 or 5, the majority of children still cannot answer this problem at the end of Grade 5, despite it being set at the Grade 3 level.

Figure 4: National School Effectiveness Study – Proportion of students answering the item correctly by Grade and quintile of student socioeconomic status (NSES Question 42)



It is important to remember that while the NSES mathematics test (set at the Grade 3 level) was the same in Grades 3, 4 and 5, the expectations of the curriculum in each year proceeded unhindered by the fact that most children still had not acquired the necessary foundational skills in the previous Grade. Weak assessment practices combined with low expectations and institutional inertia mean that most students are promoted to the next grade irrespective of whether or not they have acquired the necessary skills in the previous grade (Van der Berg et al., 2011). The growing disconnect between the real mathematics proficiency of students relative to the expectations of the curriculum mean that students fall further and further behind even while they proceed to higher Grades eventually leading to a situation of “silent exclusion” (Lewin, 2009).

5. Moving from learning deficits to learning trajectories

While the previous sections have identified the proportion of students that are not operating at a Grade 3 level, they do not provide much guidance in terms of learning trajectories into later Grades. The figures above show that some students are only learning part of the Grade 3 curriculum in either Grade 4 or Grade 5 and that many never seem to acquire these skills. However one cannot say to what extent they are also acquiring Grade 4 level skills in Grade 4 and Grade 5 level skills in Grade 5, although this is unlikely. This is because the NSES test was set at a Grade 3 level with only a small number of questions set at the Grade 4 level. One could use SACMEQ (Grade 6) and TIMSS (Grade 9) as measures of mathematical proficiency at higher levels, but these tests are not calibrated to be comparable to each other, or to earlier tests like the NSES. This is problematic since learning trajectories require data points distributed across the full range of educational phases which are comparable to each other both in terms of the content tested and the difficulty level of the tests. One alternative method to partially overcome the lack of *inter-survey* comparability is to measure the size of learning deficits in each data set using *intra-survey* benchmarks.

While most benchmarks in education are norm-referenced benchmarks (like being able to read by the age of eight), it is also possible to use the achievement level of an identifiable group as one benchmark, particularly when the composition of that group is relatively stable over time. For the purposes of the present analysis we create a benchmark which is equal to the average performance of South Africa’s quintile five students (i.e. the wealthiest 20% of students based on student socio-economic status) in each survey. There are three reasons why we believe this is a useful and appropriate benchmark: (1) Given the low intra-generational social mobility in South Africa, there is a strong case to be made that the size and composition of the wealthiest 20% of students is relatively stable over time; (2) Previous South African research has shown that this particular grouping of students performs noticeably better than the South African average, and can be seen as having its own data generating process (Spaull, 2013); and (3) The quintile system is a

widely used and recognized form of classification appearing in government reports and academic research alike.

We calculate the average performance of quintile five students for each of the following three assessments: NSES 2007/8/9 for Grades 3, 4 and 5; SACMEQ 2007 for Grade 6; and TIMSS 2011 for Grade 9. To limit the effect of possible confounding language factors, we only use the sub-set of 15 no-language items for the NSES Grade 3, 4 and 5 scores. We use the average performance of quintile five students as the reference category and compare other levels of performance (quintile and province) to these within-survey benchmarks. We set the Quintile 5 average to be equal to the “grade-appropriate level” and compare all other levels of performance relative to this Quintile 5 average. It is important to note that this is necessarily a lower-bound estimate of curriculum mastery or grade-appropriate performance since some Quintile 5 students will not be performing at the grade appropriate level. The preceding analysis of the Systemic Evaluation 2007 and NSES 2007/8/9 has shown that this is in fact the case – many Quintile 5 students are performing well below the expectations of the curriculum. Notwithstanding the above, we still believe this is a useful benchmark against which to compare other sub-groups. While the ultimate aim of any education system is to ensure that all children attain the full curriculum and exhibit sufficient mastery of it, we take the position that comparisons to the tangible group of Quintile 5 students in the country has more conceptual purchase than pegging the benchmark to a somewhat arbitrary point of curriculum mastery that is in any event not possible to do with the current data.

By using all three data sets (NSES, SACMEQ and TIMSS), we are able to calculate the difference in scores between the average Quintile 5 student and the average student in a particular sub-group, say Quintile 1 (poorest 20% of students). However, given that each of the three surveys uses a different metric to measure student performance it is not possible to use raw survey-specific scores to make comparisons across grades. To overcome this comparability problem we use the within-survey national standard deviation of South Africa as a unit of measurement. Given that the standard deviation is not a function of the specific unit of measurement (like SACMEQ points or TIMSS points) but rather a statistic describing the distribution of performance, it is possible to compare differences in student achievement across surveys that are otherwise not comparable.

One can go further and convert these standard deviation differences into Grade-level differences, as has been done in other countries. Using seven nationally normed tests of student reading and mathematics achievement, Hill *et al.* (2007, p. 172) compare the annual learning gain per Grade for American students from Grade K – 12 in standard deviations. They find that the annual learning gains vary by Grade with greater gains at earlier Grades. For example, in mathematics the learning from Grade 1 to 2 was 1.03 standard deviations, from Grade 4 to 5 was 0.56 standard deviations and from Grade 8 to 9 was 0.22 standard deviations (Hill et al., 2007, p. 172). The average math

test score gain across all seven Grade levels was 0.47 standard deviations per year, which has been used elsewhere as a benchmark for one Grade-level of learning in America (Washington State Institute for Public Policy, 2011). Unfortunately, similarly rich data does not exist for South Africa. The only two data sets which allow for the estimation of learning gains in South Africa are the NSES study (2007/8/9) for primary school, and TIMSS (2003) study for high school. Given that NSES followed the same students over time as they moved from Grade 3 into Grade 4 and 5 and tested these students using the same test, one can estimate the amount of learning between Grade 3 and 4 as a percentage of the average standard deviation between the two years. One can also calculate the learning gains between Grade 4 and 5 using NSES although these are likely to be biased given that the NSES test was set at the Grade 3 level.

When using the NSES numeracy tests to calculate learning gains there are two important caveats: firstly, one should use only those items that have no language content in them to ensure that the gains are due to increased numeracy proficiency rather than increased proficiency in the English language (as discussed above), and secondly, the results of the analysis are likely to be different based on whether one uses the panel sample (i.e. only those we can follow across all years), or the full sample (i.e. all students in each Grade). Table 3 below reports the average numeracy score for Grade 3, 4 and 5 as well as the learning gains (both in percentage points and as a percentage of the average standard deviation between the two years) for both the full numeracy test and the sub-set of 15 no-language items. As a robustness check we also impute¹³ scores for those Grade 3 children who we cannot find in the Grade 4 and Grade 5 NSES sample either due to dropout, moving or Grade repetition.

¹³ The predicted scores were calculated by first regressing the 2008 numeracy scores on the 2007 numeracy scores and including other explanatory variables such as a student's gender, socio-economic status, whether the student is over age or too old, whether a student's home language is English, whether the student is part of a large household as well as school fixed effects. This regression included only those students who were observed in both 2007 and 2008, and the coefficients were then used to predict the 2008 scores of those students who were not observed in 2008. The resultant scores were imputed into the 2008 numeracy distribution to render a new numeracy score distribution with which to do the sensitivity analysis. The process was repeated for the 2009 numeracy scores, regressing the 2009 scores on the new 2008 numeracy scores, including the imputed values.

Table 2: Quantification of a year's worth of learning in South Africa (NSES and TIMSS)

NSES			Percentage Points				Gains as % SD
			Grade 3	Grade 4	Grade 5	Gains	
Full numeracy test	Panel sample	Gr3 - Gr4	29.38	35.5		6.12	0.28
		Gr4 - Gr5		35.5	47.04	11.54	0.54
	Full sample	Gr3 - Gr4	20.22	27.88		7.66	0.38
		Gr4 - Gr5		27.88	39.66	11.78	0.58
	Full sample – imputed	Gr3 - Gr4	25.85	32.33		6.48	0.31
		Gr4 - Gr5		32.33	44.16	11.84	0.61
Sub-set of 15 no-language items only	Panel sample	Gr3 - Gr4	27.64	34.53		6.88	0.28
		Gr4 - Gr5		34.53	40.99	6.46	0.43
	Full sample	Gr3 - Gr4	17.92	26.47		8.54	0.38
		Gr4 - Gr5		26.47	38.84	12.38	0.52
	Full sample – imputed	Gr3 - Gr4	23.78	31.43		7.65	0.33
		Gr4 - Gr5		31.43	42.81	11.37	0.52
TIMSS			Standardized TIMSS Points			Gains as % SD	
			Grade 8	Grade 9	Gains		
	Full sample	Gr8 - Gr9	264	285		21	0.2

While the NSES is helpful to estimate the amount of learning per Grade in Grade 3, the best data set to calculate the amount of learning at a higher Grade is TIMSS 2003. In TIMSS 2003 the principal investigators decided to test both Grade 8 students and Grade 9 students (from the same sampled schools) using the same Grade 8 test. This was out of a concern that the TIMSS Grade 8 test was too difficult for South African Grade 8 students and thus that future administrations of TIMSS may be done at the Grade 9 level and would need a baseline for comparability. This is in fact what happened in TIMSS 2011 when only Grade 9 students were tested using the Grade 8 test. By comparing the average TIMSS score of Grade 8 and 9 students (from the same schools) on the same test and calculating this as a percentage of the South African TIMSS 2003 standard deviation, one can get an estimate of the amount of learning in one Grade at the high school level. While it would be ideal to follow the same students from Grade 8 to Grade 9 (as NSES did between Grades 3, 4 and 5), this has not been done before in South Africa and thus the best estimate available is that of the TIMSS 2003 Grade 8 and 9 students from the same schools on the same test.

One other method of calculating Grade-level equivalents is to use the benchmarks calculated by cross-national testing regimes themselves. For example, the Trends in International Mathematics and Science (TIMSS) study estimates that within a 4-year testing cycle a country could improve by a maximum of 40 points which is referred to elsewhere as “one Grade level” (Reddy, et al., 2012, p. 3). This is equal to 0.4 TIMSS standard deviations and 0.5 South African TIMSS standard

deviations.¹⁴ While this is a useful measure for comparing improvements across countries, it has not been calibrated using South African data and is therefore not specific to South Africa but rather a generic loose measure for cross-country comparisons. As we have shown in Table 3 above, the real level of learning occurring between Grade 8 and Grade 9 in South Africa is only 20 points (0.2 TIMSS 2003 South Africa standard deviations).

Since there are numerous estimates for “learning gains” presented in Table 3, it is important to motivate for the particular learning gain estimates we will use for the remainder of the paper. Given that the test was calibrated at the Grade 3 level, the distribution of the Grade 5 students on the Grade 3 test may not be an accurate reflection of the true Grade 5 distribution since it may be constrained due to a ceiling-effect leading to over-concentration at the top end of the distribution. Consequently, it is arguable that the learning gains between Grades 3 and 4 are a more accurate reflection of true learning gains than those between Grade 4 and 5 in NSES¹⁵. Secondly, given that we are only trying to measure the increase in mathematical proficiency and not the portion attributable to increased language competency, it is arguable that the estimates using the sub-set of no-language items is more accurate than those for the full test. Furthermore, if one uses the full test results for Grade 3 NSES, it will necessarily overestimate the learning between Grades 3 and 4 due to underestimating the baseline learning in Grade 3. Lastly, if one has to choose between the full sample and the panel sample (i.e. only those we can follow from Grades 3 to 4), we believe that when trying to estimate learning in a year it makes sense to choose the panel sample. This is because the students who are in the NSES Grade 4 sample but who are not in the Grade 3 sample are more likely to have repeated Grade 4 and thus overestimate the amount of learning occurring in Grade 4. As a result of the above we decide to use the no-language balanced panel estimate for the learning gain for a single year between Grades 3 and 4, i.e. 0.28 standard deviations. Incidentally this is the same as the learning gain seen in the full test balanced panel sample for the same Grades. For the learning gain between Grade 8 and Grade 9 there is only one estimate: 0.2 standard deviations (using TIMSS). Given that all of these tests were administered at the end of the year, the learning gains are for the later Grade, i.e. 0.28 is the learning that occurs in Grade 4 and 0.2 is the learning gain that occurs in Grade 9, on average, in South Africa.

Given that there are in essence only two points in the South African system for which we have psychometrically comparable data for a year of learning (Grades 3 to 4 and Grades 8 to 9), and also due to a lack of South African scholarship in this area with which to compare the above

¹⁴ The TIMSS standard deviation is roughly 100 points while the South African TIMSS 2011 standard deviation was 86 points (Mullis, et al., 2012, p. 488).

¹⁵ Looking only at grades 3, 4 and 5 in Figure 6 one may be tempted to conclude that students in quintiles 1-4 are catching up to those in quintile 5, however, a large part of this explanation is the aforementioned ceiling effect where quintile 5 students cannot score higher than 100% on the test, even while quintile 1-4 students achieve higher marks in grades 4 and 5. It is for this reason that we focus on the change between grade 3 and 4 rather than that between 4 and 5. Also see the online technical appendix where the exact scores for each quintile are reported.

results, we are sceptical of defining Grade-specific learning gains as do Hill et al. (2007). Instead we opt for a single rough estimate used uniformly across the Grades. Given the estimates presented in Table 3 above and the preceding motivation we believe that a reasonable rule of thumb for a year of learning in South Africa is 0.3 standard deviations. As a sensitivity analysis the learning trajectories were also calculated using 0.2 and 0.4 standard deviations for a year of learning. This only changes the size of the gap in predictable ways with the gap being larger the smaller the yearly learning gains¹⁶. From this it is evident that the choice of learning gain does make a difference.

Applying the above method we calculate the difference in average achievement between Quintiles 1 (poorest 20% of students) and quintile 5 (wealthiest 20% of students) for the different surveys and then convert these into a common standard-deviation metric. The difference between quintiles 1 and 5 is 28 percentage points in NSES Grade 3, 130 SACMEQ points in Grade 6, and 122 TIMSS points in Grade 9. These different metrics are not directly comparable and there is no simple way of equating the scores. Consequently we convert the differences into within-survey standard deviations and then, using the 0.3 standard deviation benchmark as one year of learning, one can say that this difference was equal to 4 Grade-levels in Grade 3¹⁷ (NSES), 4.4 Grade-levels in Grade 6 (SACMEQ) and 4.7 Grade-levels in Grade 9 (TIMSS).

Lewin (2007) provides a useful conceptual model for the trajectory needed to reach a particular goal – in this case matric (Grade 12). He refers to an ‘on-track-line’ and an ‘off-track-line’ where the off-track-line is any line below the on-track-line. In the present example, the on-track-line is calibrated to be equal to the average performance of Quintile 5 students.

To illustrate the above in a graph, we set the average Quintile 5 achievement to be equal to the Grade-appropriate benchmark such that the learning trajectory of these students are on the “on-track” trajectory and will reach matric (Grade 12) performing at roughly a Grade 12 level. We then calculate the difference between this ‘benchmark performance’ and the average performance of Quintiles 1, 2, 3 and 4 and then convert this difference into Grade-level equivalents using 0.3 standard deviations as equal to one Grade-level of learning. In doing so, we essentially create a learning trajectory spanning from Grade 3 (NSES) to Grade 9 (TIMSS) with linear projections for those Grades where we do not have data (Grade 7, 8, 10, 11 and 12). The exact figures for all calculations are provided in the online appendix.

Figure 6 below shows the likely learning trajectories of the average student in each quintile of student socioeconomic status.

¹⁶ From the results provided in the online appendix it is possible to calculate the size of these gaps for a given level of yearly-learning.

¹⁷ This would essentially place most Quintile 1 Grade 3 students at a level below Grade 1, i.e. Grade R.

Figure 6 shows that the average student in Quintile 1, 2 and 3 is functioning at approximately three Grade-levels lower than the Quintile 5 benchmark in Grades 3, 4, 5 and 6. Observing average performance by quintile in Grade 9 shows that the difference between Quintile 1, 2 and 3 students and Quintile 5 students (the benchmark) has now grown to more than four Grade-levels. If it is assumed that Quintile 5 students in Grade 9 are functioning at roughly a Grade 9 level, then Quintile 1 and 2 students are functioning at roughly a Grade 4.5 level in Grade 9. The trajectory lines, one for Quintile 5 and one for the average of Quintiles 1-4, show that in Grade 3 there already exist large differences in performance (approximately three Grade-levels) and that by the time children enter Grade 9 this gap in performance has grown to about four Grade-levels. The linear trend in performance between these two groups suggests that if the same number of students in Quintiles 1-4 in Grade 9 continued in schooling until Grade 12 (i.e. no drop out between these two periods) they would be functioning at approximately 4.9 Grade levels lower than their Quintile 5 counterparts in Grade 12 (1.5 standard deviations lower).

The reason why one cannot easily use the matric (Grade 12) data as another point in the learning trajectory is the substantial number of students that drop out of schooling between Grade 9 and Grade 12 in South Africa. Taylor (2012, p. 6) shows that the average enrolment in Grades 4, 7 and 10 between 2008 and 2011 in South Africa was approximately 1,000,000 in each Grade, but by Grade 12 this figure drops to roughly 600,000 students. Consequently, if we were to include Grade 12 as a data point we would need to make a number of assumptions about dropout and the differential distribution of dropout across the socioeconomic spectrum. For the purposes of this paper we do not extend the analysis to Grade 12 by using matric data.

Returning to Lewin's (2007) notion of an "on-track" progress line, perhaps the most important conclusion arising from this conceptual framework is that any performance below the "on-track" line creates an increasing gradient of expectation as the pupil moves into higher grades. This expectation is what is required by the curriculum to reach the goal (passing the grade 12 exam, for example) relative to where the student is at the present. As students' learning deficits grow, the gradient of what needs to be achieved to reach the goal then progressively steepens to the point where it enters what Lewin (2007, p. 7) refers to as a 'Zone of Improbable Progress.' For example, the improvement that is required to bring the average Grade 9 Quintile 1 student in South Africa up to the required benchmark by Grade 12 is unrealistic given that they are performing at roughly a Grade 5 level in Grade 9. By contrast, the gradient of achievement required to bring the average Quintile 1 Grade 3 pupil up to the required benchmark by matric is slightly more manageable. The clear conclusion arising from this analysis is that intervening early to correct and prevent learning

deficits is the only sustainable approach to raising average achievement in under-performing schools.

What we would add to this conclusion is that the root cause of these weak educational outcomes is that children are acquiring debilitating learning deficits early on in their schooling careers and that these remain with them as they progress through school. Because they do not master elementary numeracy and literacy skills in the foundation and intermediate phases, they are precluded from further learning and engaging fully with the Grade-appropriate curriculum, in spite of being enrolled in school. Lewin (2007, p. 10) refers to these children as 'silently excluded' since they are enrolled and attending school but learning little. Importantly, these children are precluded from further learning, not because of any inherent deficiency in their abilities or aptitudes, but rather because of the systematic and widespread failure of the South African education system to offer these students sustained and meaningful learning opportunities. Indeed, many children from poorer backgrounds have both the ability and the desire to succeed, and when provided with meaningful learning and remediation opportunities, do in fact succeed (see Spaul et al, 2012 for an example).

5.1 Methodological caveat: Test comparability

Given that each of the three tests used in this analysis was developed and administered by a different organization, it is useful to provide some indication of how these tests were developed, the content that they covered and whether or not they were aligned to the South African curriculum at each grade. Full discussions of the psychometric properties of the items in each test are beyond the scope of this study but are available for each test; NSES (Taylor et al, 2013, Ch. 2), SACMEQ (Ross et al, 2005, Ch. 2) and TIMSS (Foy, et al., 2013).

The NSES numeracy test was constructed to be completely aligned with the National Curriculum Statement, which was the curriculum at the time. As mentioned previously, the test was the same as the grade 3 Systemic Evaluation of 2007, which was commissioned by the Department of Basic Education to monitor grade 3 outcomes relative to the grade 3 curriculum (Taylor et al., 2013). Approximately 60% of the items in the test covered four tasks which forms the fundamental building blocks of mathematics namely: counting and ordering whole numbers, addition, multiplication and subtraction (Taylor et al. 2013). The remainder of the problems were split between items dealing with fractions, decimals, patterns, graphs, shapes and measurement (Taylor & Taylor, 2013). The difficulty level of these questions ranged from a Grade 1 level to a Grade 4 level, as discussed in the data section above. Given that the same test was administered in grades three, four and five, one can think of the test becoming easier over time as students acquire new skills and find the test questions from earlier grades easier to understand and answer correctly. Since the NSES test was predominantly a grade 3 test, we do not interpret the learning

gains from grade 4 to grade 5 as being authoritative and prefer to use the gains between grade 3 and grade 4. This is discussed in more detail below with reference to Table 1.

The construction of the SACMEQ test was done so as to ensure congruence with the curricula, syllabi, exams and textbooks used in all of the participating countries (Ross et al., 2005). The content of the SACMEQ test falls under three broad domains namely number, space and data, and measurement. Given that there are multiple countries that participate in SACMEQ, and that the SACMEQ assessments need to find common domains across most education systems, these tests can be thought of as assessing the core mathematics curriculum and competencies at the grade 6 level (Ross et al, 2005). In the South African SACMEQ 2007 report, written by the South African Department of Basic Education, they explain that “In the national curriculum statement emphasis is placed on teachers designing tasks in such a way as to ensure that a variety of skills are assessed. The eight SACMEQ levels for reading literacy and mathematics presented in this report provide an *appropriate benchmark* to model assessments and structure learning such that learners may be exposed to the expected range of competencies for their age group (Moloi & Chetty 2011, p. 7; emphasis added).

The TIMSS mathematics test covered the broad content areas of number, data and chance, algebra and geometry, and the cognitive domains of reasoning, knowing and applying (Mullis et al., 2012). A comparison between the TIMSS 2011 mathematics assessment framework and the Revised National Curriculum Statement (the curriculum in use at the time of testing) indicates that there is a 94% overlap (Reddy, et al., 2012). It is also important to remember that South Africa takes part in TIMSS by testing its grade 9 students despite this being a grade 8 test internationally.

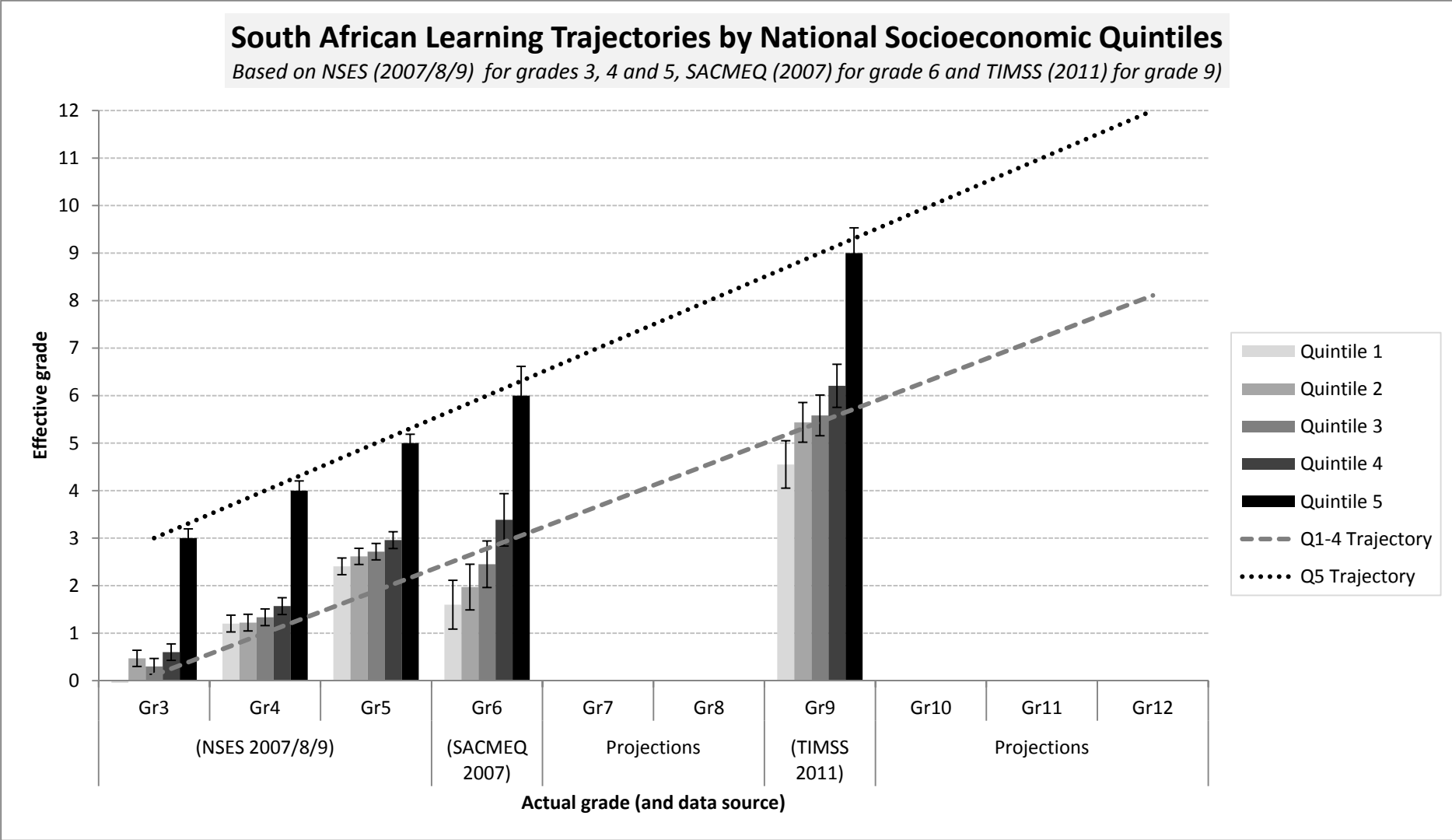
As can be seen in the discussion above, the type of mathematics tested in each of these tests differs to some degree between the three assessments since each test may place more or less weight on a particular learning area. This is an important point since it is possible that student outcomes (or the gaps between rich and poor students) is also a function of the items on the test rather than their true performance (or the true gaps between rich and poor students). For example, if one looks at TIMSS, South African students achieved at the bottom of the international TIMSS 2011 mathematics rankings, with 32% performing no better than random guessing (Mullis et al., 2012, p. 457). Consequently, it is prudent to ask whether or not the 2011 TIMSS international grade 8 test was more challenging than the grade 9 curriculum in South Africa. If this is the case – and the performance of quintile 5 students declines less than that of students in quintile 1-4 as a result - then the gap between rich and poor could be seen to grow between grade 6 (SACMEQ) and grade 9 (TIMSS) when perhaps the gap remained unchanged in reality.

However this does not seem to be the case. If one looks at the performance of South African grade 9 students on the TIMSS 2011 grade 8 mathematics test, one can see that only 3% of students

achieved the 'High' or 'Advanced' TIMSS benchmarks (Reddy et al., 2012, p. 11). If one compares this to the performance of all grade 9 students on the South African grade 9 Annual National Assessments (ANA) conducted in 2012 and 2013 in South Africa, one sees similar results. Only 2-3% of grade 9 students in each year reached "Acceptable achievement" as defined by the Department of Basic Education (DBE, 2013, p. 53). Importantly these tests are specifically aligned to the South African curriculum. Given these low results the Minister of Basic Education in South Africa convened a task team to look at the grade 9 mathematics ANA test to determine if it was too difficult. The task team concluded that the test was "fair, valid and reliable" leading the Minister to conclude: "the results are a genuine and credible reflection of the learning achievements in grade 9 maths" (Motshekga, 2013). Therefore, while it is true that South African students do seem to find the TIMSS test more difficult, this is largely because they are falling behind relative to the curriculum not because the tests are unreasonably difficult relative to the curriculum.

To summarize the methodological discussion above, it has been argued that the three tests (NSES, SACMEQ and TIMSS) are a relatively accurate representation of the broad mathematics achievement of South African students at each stage (grade 3, 4, 5; 6 and 9 respectively). The aim in using these three assessments is not to estimate the gaps in learning with pinpoint precision – that would require longitudinal data. However, longitudinal data spanning these seven years is not available in South Africa. Consequently, we use multiple cross-sectional datasets and argue that they are broadly aligned to the curriculum at grade 3, 6 and 9. It is possible that this curriculum-alignment assumption is false. If, for example the grade 9 test (TIMSS) is more difficult than the average mathematics found in the curriculum at grade 9, the gap between quintile 5 and quintile 1 could possibly increase even if the 'true' gap remains unchanged. This would only be the case if the standard deviation did not increase and simultaneously quintile 1-4 students did disproportionately worse than quintile 5 students. However it is also possible that a more right-skewed distribution (due to a more difficult test) could decrease the standard deviation due to additional bunching at the bottom of the distribution. Given that the gap in years is a function of both the standard deviation and the absolute gap between the quintile 5 and the other quintiles (and that these could move in different directions), it is unclear what the net-effect would be on the size of the gap if the tests were of vastly differing curriculum-alignment. However, as is argued above, we do not believe that any of these tests is grossly misaligned with the curriculum at that grade.

Figure 6: South African Mathematics Learning Trajectories by national socioeconomic quintiles (Based on NSES 2007/8/9 for Grades 3/4/5, SACMEQ 2007 for Grade 6 and TIMSS 2011 for Grade 9, including 95% confidence interval)



6. Conclusion

The above analysis has provided an overview of the size and distribution of learning deficits in the South African education system. Using local and international assessments of mathematics achievement and converting test-score gaps into standard deviations and then into Grade-levels of learning, it was possible to estimate empirically and illustrate graphically the learning trajectories of wealthy and poor students in South Africa. The key finding emerging from this research is that by Grade 3, children in Quintiles 1-3 are already three years' worth of learning behind their Quintile 5 peers and that this gap grows as they progress through school to the extent that by Grade 9 they are four years' worth of learning behind their Quintile 5 peers. Previous studies have shown that this low quality of education offered to the poor eventually becomes a poverty trap (Van der Berg, et al., 2011). Thus one can say that poor children in South Africa, who make up the majority, are starting behind and staying behind, casting doubt on the ability of the South African schooling system to impart to students the knowledge, skills and values they need to become full members of society and thus promote social mobility.

The clear policy recommendation which proceeds from these findings confirms what is becoming increasingly accepted, that any intervention to improve learning in South Africa needs to intervene as early as possible. Given South Africa's egregiously high levels of inequality, it should come as no surprise that poor children in South Africa find themselves at a nexus of disadvantage, experiencing a lack of social, emotional and cognitive stimulation in early childhood. These children then enter a primary school system that is unable to equip them with the skills needed to succeed in life, let alone to remediate the large learning deficits they have already accumulated to date.

In this regard the South African government has begun to universalize a pre-Grade-1 year of school termed 'Grade-R.' Between 2001 and 2012 the number of Grade R places in South African schools (public and independent) increased threefold from 242,000 to 768,000 leading to a situation where 78% of 5-year olds were in some sort of education programme in 2009 (Van der Berg et al., 2013. p.3). However, the research of Van der Berg et al. (2013) has shown that there are serious deficiencies in the quality of this additional year of pre-school education. They find that "...the impact of Grade R in South Africa is small and there is virtually no measurable impact for the poorest three school quintiles, while there are some impact for the higher quintile schools...quality thus needs attention" (Van der Berg et al., 2013, p.3). This indicates that the emphasis on access to Grade R alone has not assisted in bridging the learning gaps between rich and poor, and that more attention should be paid to the quality of the Grade R provided. This extends to the quality of teachers employed, the training and support they are provided and a curriculum that clearly guides teachers to understand how children learn at this age.

When faced with limited resources and a choice of where to intervene in the schooling system, the counsel from both the local and international literatures is unequivocal; the earlier the better. The need to focus on the primary Grades, and especially the pre-primary years, is not only driven by the fact that underperformance is so widespread in these phases, but also because remediation is most possible and most cost-effective when children are still young (Heckman, 2000). Due to the cumulative negative effects of learning deficits - particularly for vertically-integrated subjects like mathematics - it is not usually possible to fully remediate pupils if the intervention is too late (i.e. in high school), as too many South African interventions are. Nobel Laureate Professor James Heckman summarises the above succinctly when he explains that:

“Policies that seek to remedy deficits incurred in early years are much more costly than early investments wisely made, and do not restore lost capacities even when large costs are incurred. The later in life we attempt to repair early deficits, the costlier the remediation becomes” (Heckman, 2000, p. 5).

7. References

- Aubrey, C., Dahl, S. & Godfrey, R., 2006. Early Mathematics Development and Later Achievement: Further Evidence. *Mathematics Education Research Journal*, 18(1), pp. 27-46.
- Aubrey, C. & Godfrey, R., 2003. The Development of Children's Early Numeracy through Key Stage 1. *British Educational Research Journal*, 29(6), pp. 821-840.
- Aunio, P. & Niemivirta, M., 2010. Predicting children's mathematical performance in grade one by early numeracy. *Learning and Individual Differences*, pp. 427-435.
- Aunola, K., Leskinen, E., Lerkkanen, M. & Nurmi, J., 2004. Developmental Dynamics of Math Performance From Preschool to Grade 2. *Journal of Educational Psychology*, 96(4), pp. 699-713.
- Banerjee, A. V. & Duflo, E., 2011. Top of the class. In: *Poor Economics: A radical rethinking of the way to fight global poverty*. New York: Public Affairs, pp. 71-102.
- Carnoy, M., Chisholm, L., Addy, N., Arends, F., Baloyi, H., Irving, M., et al. 2011. *The Process of Learning in South Africa: The Quality of Mathematics Teaching in North West Province*, Cape Town: HSRC Press.
- DoE , 2002. *Revised National Curriculum Statement Grades R-9. Mathematics*, Pretoria: Gazette No. 23406. Vol 443.
- Desoete, A. et al., 2009. Classification, Seriation, and Counting in Grades 1, 2 and 3 as Two-Year Longitudinal Predictors for Low Achieving in Numerical Facility and Arithmetical Achievement. *Journal of Psychoeducational Assessment.*, 27(3), pp. 252 - 264.
- DoE, 2008b. *Ministerial Committee on Learner Retention in the South African Schooling System report*, Pretoria: Department of Education.
- DoE, 2008. *Grade 3 Systemic Evaluation 2007 Leaflet*, Pretoria: Department of Education.
- DoBE, 2013. Report on the Annual National Assessment of 2013: Grade 1 to 6 & 9. Pretoria. Department of Basic Education
- Ensor, P. et al., 2009. Specialising pedagogic text and time in Foundation Phase numeracy classrooms. *Journal of Education*, Issue 47, pp. 5-30.
- Fiske, E. & Ladd, H., 2004. *Elusive Equity: Education Reform in Post-apartheid South Africa*. Washington: Brookings Institution Press / HSRC Press.
- Fleisch, B., 2008. *Primary Education in Crisis: Why South African schoolchildren underachieve in reading and mathematics*. Cape Town: Juta & Co.
- Foy, P., Arora, A. & Stanco, G. M., 2013. *TIMSS 2011 User Guide for the international database*, Boston: TIMSS & PIRLS Study Center.
- Foy, P., Martin, M. & Mullis, I., 2010. *The limits of measurement: problems with estimating reading achievement in PIRLS 2006 for low-performing countries*, Boston: TIMSS & PIRLS International Study Center.

- Gagne, R., 1962. The acquisition of knowledge. *Psychological Review*, Volume 69, pp. 355-365.
- Heckman, J., 2000. Policies to foster human capital. *Research in Economics*, Issue 54, pp. 3-56.
- Hill, C., Bloom, H., Black, A. & Lipsey, M., 2007. *Empirical benchmarks for interpreting effect sizes in research*, MDRC Working Papers on Research Methodology.
- Howie, S. & Hughes, C., 1998. *Mathematics and science literacy of final-year school students in South Africa: a report on the performance of South African students in the Third International Mathematics and Science Study (TIMSS)*. Pretoria: Human Sciences Research Council.
- Hungi, N. et al., 2010. *SACMEQ III Project Results: Pupil Achievement Levels in Reading and Mathematics*, Paris: Southern and Eastern Africa Consortium for Monitoring Educational Quality.
- Jordan, N., Kaplan, D., Locuniak, M. & Ramineni, C., 2007. Predicting First-Grade Math Achievement from Developmental Number Sense Trajectories. *Learning Disabilities Research & Practice*, 22(1), pp. 36-46.
- Kilpatrick, J., Swafford, J. & Findell, B., 2001. *Adding It Up: Helping Children Learn Mathematics*. Washington, DC: National Academy Press.
- Lewin, K., 2007. *Improving Access, Equity and Transitions in Education: Creating a Research Agenda*, Sussex: Consortium for research on Educational Access, Transitions and Equity (CREATE).
- Lewin, K. M., 2009. Access to Education in Sub-Saharan Africa: patterns, problems and possibilities. *Comparative Education*, 45(2), pp. 151-174.
- Moloi, M. & Chetty, M., 2011. *The SACMEQ III Project in South Africa: A Study of the Conditions of Schooling and the Quality of Education*, Pretoria: Department of Basic Education.
- Moloi, M. & Strauss, J., 2005. *The SACMEQ II Project in South Africa: A Study of the Conditions of Schooling and the Quality of Education*, Harare: SACMEQ.
- Motshekga, A. 2013. Statement on the Release of Annual National Assessments Results for 2013. Delivered at Mahlahle Primary School, Atteridgeville.
- Mullis, I., Martin, M., Foy, P. & Arora, A., 2012. *TIMSS 2011 International Results in Mathematics*, Chestnut Hill, USA: TIMSS & PIRLS International Study Center and IEA.
- Muralidharan, K. & Zieleniak, Y., 2013. *Measuring Learning Trajectories in Developing Countries with Longitudinal Data and Item Response Theory*. Oxford.
- Murimba, 2005. The Southern and Eastern African Consortium for Monitoring Educational Quality (SACMEQ): Mission, approach and projects. *Prospects*, XXXV(1).
- Posner, G. & Strike, K., 1976. A categorization scheme for principle of sequencing content. *Review of Education Research*, 46(4), pp. 665-690.
- Pritchett, L. & Beatty, A., 2012. The negative consequences of overambitious curricula in developing countries. *HKS Faculty Research Working Paper Series RWP12-035, John F. Kennedy School of Government, Harvard University*.

Reddy, V., 2006. *Mathematics and Science Achievement at South African Schools in TIMSS 2003*, Cape Town: HSRC Press.

Reddy, V. et al., 2012. *Highlights from TIMSS 2011: The South African perspective*, Pretoria: HSRC.

Reeves, C. & McAuliffe, 2012. Is curricular incoherence slowing down the pace of school mathematics in South Africa? A methodology for assessing coherence in the implemented curriculum and some implications for teacher education. *Journal of Education*, 61(5), pp. 9-36.

Reeves, C. & Muller, J., 2005. Picking up the pace: variation in the structure and organization of learning school mathematics. *Journal of Education*, Volume 37.

Ross, K. et al., 2005. The Conduct of the SACMEQ III Project. In: *The SACMEQ II Project in Kenya: A Study of the Conditions of Schooling and the Quality of Education*. Harare: SACMEQ.

Scandura, J. M. & Wells, J. N., 1967. Advance organizers in learning abstract mathematics. *American Education Research Journal*, pp. 295-301.

Schollar, E., 2008. *Final Report (Short Version) The Primary Mathematics Research Project 2004-2007: Towards evidence-based educational development in South Africa*. Johannesburg: Eric Schollar and Associates.

Spaull, N., Burger, R., Burger, C., Van Wyk, C., Van der Berg, S., & Dzivakwi, R. 2012. Against the Odds: An Evaluation of the Ikamva Youth Programme. An Evaluation Report for Ikamva Youth and the DG Murray Trust. Department of Economics, Stellenbosch University.

Spaull, N., 2013. Poverty & privilege: Primary school inequality in South Africa. *International Journal of Educational Development*, Volume 33, pp. 436-447.

Spaull, N., 2013. *South Africa's Education Crisis*, Johannesburg: Center for Development and Enterprise (CDE).

Taylor, N., 2002. *Accountability and Support: Improving Public Schooling in South Africa*, Johannesburg: JET Education.

Taylor, N., Muller, J. & Vinjevold, P., 2003. *Getting Schools Working*. Cape Town: Pearson Education.

Taylor, N. & Reddi, B., 2013. Writing and learning mathematics. In: *Creating Effective Schools*. Cape Town: Pearson.

Taylor, N., Van der Berg, S. & Mabogoane, T., 2013. *What makes schools effective? Report of the National Schools Effectiveness Study*. Cape Town: Pearson.

Taylor, N. & Vinjevold, P., 1999. *Getting Learning Right: Report of the President's Education Initiative Research Project*, Johannesburg: Joint Education Trust.

Taylor, S., 2012. *A Note on Matric Result Trends*, Unpublished report.

Taylor, S. & Coetzee, M., 2013. Estimating the impact of language instruction in South African primary schools: a fixed effects approach. *Stellenbosch Economic Working Papers: 21/13*, pp. 1-33.

- Taylor, S. & Taylor, N., 2013. Learner Performance in the NSES. In: *Creating Effective Schools*. Cape Town: Pearson.
- Taylor, S., Van der Berg, S., Reddy, V. & Janse van Rensburg, D., 2011. How well do South African schools convert grade 8 achievement into matric outcomes? *Stellenbosch Economic Working Papers*, pp. 1-46.
- Taylor, S. & Yu, D., 2009. The Importance of Socioeconomic Status in Determining Educational Achievement in South Africa. *Stellenbosch Economic Working Papers*.
- Van der Berg, S., 2007. Apartheid's Enduring Legacy: Inequalities in Education. *Journal of African Economies*, 16(5), pp. 849-880.
- Van der Berg, S., Burger, C., Burger, R., De Vos, M., Du Rand, G., Gustafsson, M., Moses, E., Shepherd, D., Spaul, N., Taylor, S., Van Broekhuizen, H., Von Fintel, D., 2011. Low quality education as a poverty trap. Report for the Program to Support Pro-Poor Policy Development (PSPPD). Stellenbosch University.
- Van der Berg, S., Girdwood, E., Shepherd, D., Van Wyk, C., Kruger, J., Viljoen, J., Ezeobi, E. & Ntaka, P. 2013. The Impact of the Introduction of Grade R on Learning Outcomes. Research on Socio-Economic Policy (ReSEP) report for the Department of Basic Education and the Department of Performance and Monitoring Evaluation in the Presidency.
- Venkat, H. & Naidoo, D., 2012. Analyzing coherence for conceptual learning in a Grade 2 numeracy lesson. *Education as Change*, Volume 16, pp. 21-33.
- Washington State Institute for Public Policy, 2011. *The Economic Value of Learning Time in K-12 Schools: A Summary of Research Evidence and an Economic Analysis*. Washington

Online appendix

Achievement scores for various assessments (NSES using no-language items only)

		NSES 2007/8/9						SACMEQ 2007		TIMSS 2011	
		Grade 3	SE	Grade 4	SE	Grade 5	SE	Grade 6	SE	Gr9	SE
Student wealth quintiles	Quintile 1	13.74	0.039	27.25	0.039	37.44	0.044	452.80	4.14	316	5.7
	Quintile 2	24.56	0.040	27.56	0.039	40.06	0.044	463.80	3.15	318	3.6
	Quintile 3	23.49	0.040	29.09	0.041	41.10	0.047	478.00	3.42	336	4
	Quintile 4	25.93	0.046	32.30	0.045	44.39	0.052	505.70	5.14	360	5.6
	Quintile 5	43.30	0.064	48.90	0.066	59.11	0.062	583.10	6.56	438	9.7
Province averages	ECA	26.34	0.047	29.83	0.046	42.55	0.052	468.77	10.3	316	
	FST	28.26	0.072	31.78	0.069	42.48	0.068	491.57	10.1	359	
	KZN	26.31	0.047	36.72	0.045	45.35	0.048	485.23	8.22	337	
	LMP	14.60	0.040	24.31	0.045	36.27	0.053	446.72	5.3	322	
	MPU	26.23	0.074	31.76	0.071	45.08	0.074	476.12	8.2	344	
	NCA	26.46	0.142	35.88	0.134	44.82	0.158	498.72	10.8	366	
	NWP	27.71	0.077	30.36	0.076	45.54	0.086	503.00	13.1	350	
	WCA	36.74	0.089	46.85	0.081	59.85	0.079	565.69	12	404	
Province averages excluding Q5	ECA	23.83	0.046	27.65	0.043	40.30	0.050	465.14	9.8		
	FST	21.82	0.061	26.92	0.062	38.41	0.063	471.74	6.4		
	KZN	23.66	0.044	33.52	0.039	42.23	0.044	465.88	6.4		
	LMP	14.17	0.041	24.06	0.047	35.39	0.055	441.45	3.3		
	MPU	22.29	0.073	28.55	0.069	41.33	0.076	460.89	4.1		
	NCA	20.83	0.128	32.68	0.132	39.72	0.157	473.00	5		
	NWP	22.71	0.070	25.76	0.063	41.09	0.086	462.11	4.8		
	WCA	24.49	0.107	39.02	0.105	54.31	0.112	518.53	8.1		
National avg		25.85	0.023	32.74	0.022	44.34	0.024	494.84	3.8	352	2.5
National SD		22.04		20.87		21.46		98.68		86	

Number of years behind relative to Quintile 5 (benchmark)

		NSES 2007/8/9				SACMEQ		TIMSS			
		Grade 3	SE	Grade 4	SE	Grade 5	SE	Grade 6	SE	Gr9	SE
Student wealth quintiles	Quintile 1	4.7	0.011	3.4	0.012	3.4	0.012	4.4	0.262	4.7	0.436
	Quintile 2	3.0	0.011	3.4	0.012	3.0	0.012	4.0	0.246	4.7	0.401
	Quintile 3	3.1	0.011	3.1	0.012	2.9	0.012	3.6	0.250	4.0	0.407
	Quintile 4	2.8	0.012	2.6	0.013	2.3	0.013	2.6	0.282	3.0	0.434
	Quintile 5	0.0	0.014	0.0	0.015	0.0	0.014	0.0	0.313	0.0	0.532
Province averages	ECA	2.7	0.012	3.0	0.013	2.6	0.013	3.9	0.412	4.7	
	FST	2.4	0.015	2.7	0.015	2.6	0.014	3.1	0.407	3.1	
	KZN	2.7	0.012	1.9	0.013	2.2	0.012	3.3	0.355	3.9	
	LMP	4.6	0.011	3.9	0.013	3.6	0.013	4.6	0.285	4.5	
	MPU	2.7	0.015	2.7	0.015	2.2	0.015	3.6	0.355	3.6	
	NCA	2.7	0.024	2.1	0.024	2.3	0.026	2.9	0.427	2.8	
	NWP	2.5	0.015	2.9	0.016	2.2	0.017	2.7	0.495	3.4	
	WCA	1.0	0.017	0.3	0.017	-0.1	0.016	0.6	0.462	1.3	
Province averages excluding Q5	ECA	3.1	0.012	3.4	0.013	3.0	0.012	4.0	0.398		
	FST	3.4	0.013	3.5	0.014	3.3	0.014	3.8	0.310		
	KZN	3.1	0.012	2.4	0.012	2.7	0.012	4.0	0.310		
	LMP	4.6	0.012	3.9	0.013	3.8	0.013	4.8	0.248		
	MPU	3.3	0.015	3.2	0.015	2.8	0.015	4.1	0.261		
	NCA	3.6	0.022	2.6	0.024	3.1	0.026	3.7	0.279		
	NWP	3.3	0.014	3.7	0.015	2.9	0.017	4.1	0.275		
	WCA	3.0	0.019	1.6	0.020	0.8	0.020	2.2	0.352		
National avg		2.8	0.010	2.6	0.011	2.3	0.010	3.0	0.256	3.3	

Calculating effective grade by sub-group

		(NSES 2007/8/9)			(SACMEQ 2007) Projections		(TIMSS 2011) Projections					
		Gr3	Gr4	Gr5	Gr6	Gr7	Gr8	Gr9	Gr10	Gr11	Gr12	
School wealth quintiles	Quintile 1	-1.69	0.56	1.56	1.60			4.27				
	Quintile 2	0.02	0.61	1.98	1.97			4.35				
	Quintile 3	-0.14	0.86	2.14	2.45			5.05				
	Quintile 4	0.24	1.36	2.66	3.39			5.98				
	Quintile 5	3.00	4.00	5.00	6.00			9.00				
	Q1-4 Trajectory	-0.39	-0.39	-0.39	-0.39	-0.39	-0.39	-0.39	4.91	4.91	4.91	4.91
	Q5 Trajectory	3.00	4.00	5.00	6.00	7.00	8.00	9.00	10.00	11.00	12.00	
Province averages	ECA	0.31	0.97	2.37	2.14			4.27				
	FST	0.61	1.28	2.36	2.91			5.94				
	KZN	0.30	2.07	2.81	2.69			5.09				
	LMP	-1.55	0.10	1.37	1.39			4.50				
	MPU	0.29	1.28	2.77	2.39			5.36				
	NCA	0.33	1.93	2.73	3.15			6.21				
	NWP	0.52	1.06	2.85	3.29			5.59				
	WCA	1.96	3.67	5.12	5.41			7.68				
Province averages excluding Q5	ECA	-0.09	0.63	2.01	2.02							
	FST	-0.41	0.51	1.71	2.24							
	KZN	-0.12	1.56	2.32	2.04							
	LMP	-1.62	0.06	1.23	1.22							
	MPU	-0.33	0.77	2.18	1.87							
	NCA	-0.57	1.43	1.92	2.28							
	NWP	-0.27	0.33	2.14	1.91							
	WCA	0.01	2.43	4.24	3.82							