Can Workbooks Improve Learner Performance? Findings of the Randomised Control Trial of the Primary Mathematics Research Project[1]

Brahm Fleisch, Nick Taylor, Roelien du Toit, and Ingrid Sapire

Paper Presented at the Wits School of Education Research Seminar, Staff Lounge

24 August 2010

ABSTRACT

A randomised field trial was conducted in 42 Gauteng schools in quintiles 1-4, selected at random from the population of Gauteng primary schools. Two sets of materials were tested, one specifically designed as a workbook which emphasises basic skill proficiency, as a prerequisite for higher order mathematical processes. The second set of materials consisted of the standard textbook most widely used in South African schools. The trial covered 14 weeks of teaching Learning Outcome 1: numbers and operations. Both sets of learners showed significant improvement in the order of 8 percentage points off a baseline of around 40%, with the difference in gains between the groups being non-significant. However, a more detailed study of the gains in different skills areas show interesting results, which reflect differences in the respective structures of the two sets of materials.

1. Introduction

In the 2010 State of the Nation Address, the President declared that education and skills development would be at the centre of the current government's policies. He further suggested that for the 2010 programme in particular, government would focus on improving children's ability to read, write and count in the foundation years. He set specific targets, indicating that by 2014 60% of learners would achieve a pass mark (presumably meaning that learners scored above 50% on the standardised tests). To achieve this target he outlined government's strategy as follows:

> *We want learners and teachers to be in school, in class, on time, learning and teaching for seven hours a day. We will assist teachers by providing detailed daily lesson plans. To students we will provide easy-to-use workbooks in all 11 languages.*

In her speech to parliament, the Minister of Basic Education (2010) announced that an amount of R750 million had been allocated in the 2010/11 budget for these workbook reaching R1 billion in 2012/13. These workbooks thus have been identified a key component of the overall strategy to improve primary school learner performance.

Given the importance attached to workbooks in the overall strategy to improve learner performance, what is known about these learner support materials? In particular, what is the evidence base that links the provision and use of these resources to improved learning outcomes at-scale? At an operational level, what does the research show about the relative cost-effectiveness of the workbooks compared to conventional classroom textbooks and stationery?   To answer these questions, we have undertaken a randomised control trial of the workbook of the Primary Mathematics Research Project, designed to address the problems of primary mathematics teaching in South Africa.   The purpose of the study is to compare the relative gains in mathematics knowledge for two random samples of Grade 6 schools using the new workbook materials compared to a standard mathematics textbook.

This paper begins by locating this study within the debate about the relative effectiveness of different kinds of learning and teaching support materials. A review of the literature reveals very little policy-relevant research on the question of the design, use and effectiveness of written materials for the classroom. The third section of the paper provides background and a rationale for the selection of the Primary Mathematics Research Project workbook. This is followed by a discussion of the research design and methodological issues attached to the selection of a randomised control trail as the research design. The fifth section presents the major empirical findings of study. The paper concludes with a summary of the key findings and explores the research and policy implications.

2. Literature Review

Lockheed (1991) and more recently Abadzi (2006) have reviewed empirical evidence from a range of countries on the relative importance of learner materials for improving learning outcomes. Research from a wide range of contexts including Nicaragua, the Philippines, Brazil, Fiji, Ghana, and Guinea, have consistently shown improvement in learner performance when sufficient textbooks were supplied.   Advocates of textbooks argue that these resources work as they reduce wasted instructional time (Lockheed, 1991). Advocates of textbooks do acknowledge, however, that the effectiveness of these learning resources is conditional. They work if the materials are pedagogically sound, culturally appropriate and durable. They also require teachers to be trained in the use of the materials, and learners need to be able to take textbooks home.

The emphasis on provisioning of textbooks as a key part of the improvement strategy, closely associated with the World Bank approaches of the 1980s, is not without critics. Fuller (1991), for example, drawing on his field work in Botswana and Malawi, expressed scepticism about the use of textbooks in developing countries. This scepticism has recently been supported by robust empirical evidence. In a study by Glewwe, Kremer & Moulin (2007) the introduction of textbooks failed to increase overall student performance. This research showed that students who were academically strong increased their performance levels substantially with the use of textbooks, but that those students who were average or weak showed no substantial gains through textbook use.

Absent from the literature is any analysis of the relative merits of different kinds of learner materials. In particular, little is known of the relative cost-effectiveness of approach that employ workbooks, i.e. materials designed to be written directly in, compared to approaches that use conventional textbooks. The assumption made by the developers of the workbook concept is that these materials have the advantage of eliminating the need for copying questions from the board or textbook. The additional advantage is that they do not come with the normal retrieval problems associated with conventional textbooks, as they are designed to be used only once.

3. Background to the Primary Mathematics Research Project

To appraise the relative merits of workbooks as a cost-effective intervention to improve primary school achievement, the study team determined that it would be prudent to select case study materials that had been developed and field tested in South Africa and had existing empirical evidence to demonstrate effectiveness. With these criteria in mind, we reviewed possible options and selected the Primary Mathematics Research Programme (PMRP) workbooks designed and produced by Eric Schollar and Associates.

The PMRP workbooks were field tested as part of a wider intervention programme in primary mathematics (Schollar, 2008). The intervention programme, with workbooks at its centre, included the implementation of 70 mathematics lessons over 14 weeks. The assumption underlying the original PMRP workbook intervention was that mathematic performance would only be enhanced if certain 'bedrock' skills were taught systematically. These bedrock skills are associated with the capacity to perform mental calculations quickly and accurately. As Schollar notes:

> The application of these algorithms allows the solving of extremely complex calculations in simple steps through an understanding and knowledge of basic number bonds, the multiplication (and division) tables and, above all, an understanding of place value in the base-10 number system. Conversely the failure of learners to understand the number system and to master arithmetic operations beyond the reach of simple counting of single units renders learners incapable of developing any degree of mathematics proficiency. (Schollar, 2008, 19)

At the centre of the PMRP workbook intervention is an approach that stresses:

- direct instruction by the teacher;
- a hierarchically structured set of activities that begins with simple number use, rising to more complex activities such as novel problems and puzzles; and
- materials that are not based on the assessment standards for one grade but are aimed at multi-grades combined with a diagnostic test that directs learners to their initial level of competence.

The PMRP materials drew on ideas of *direct instruction* which advocate extensive drill and practice as suggested in recent developments in cognitive science (Abadzi, 2006). The materials deal only with Learning Outcome 1 (LO1, Number, Operations and relationships), as specified in the National Curriculum Statement (NCS) for Grades 4-6 mathematics. Special attention is given to the correct curriculum sequencing and progression. The learner workbooks, provided to all learners in the study, contains all content that corresponds directly with the teacher materials in terms of content, sequencing and progression. Strong emphasis is placed on providing graded exercises, and some extension.

To take account of the great variation in performance in every class, the PMRP approach begins by diagnosing which level each child is at in terms of the relevant topic: diagnostic tests are a key part of the programme. The PMRP workbook is graded and, following the diagnostic test, each child is placed at the appropriate level in the book (Cow, Lion, Elephant, Goat) which s/he then follows through the book. The lesson plans for each week, Day 1-4 have a simple logical structure, while Day 5 lesson plans prescribe review, assessment, enrichment and remediation exercises. The assumption in the workbook materials is that the prescribed lessons and exercises ensure that teaching proceeds at the correct pace required to cover the curriculum and, most important, that children work from the text, reading and writing every day.

The PMRP approach, with the workbooks at the core, was field tested in forty schools in one district in Limpopo in 2007. The schools were randomly assigned to two groups: a control group and an intervention group with 20 schools in each. The intervention took place in Grades 4 and 6 in the intervention group. Grade 4 and 6 teachers in the intervention schools attended a one and a half day training workshop at which they receive lesson plans for 14 weeks, 14 week learner workbook and related assessment materials. The control group received no intervention, and in these schools teachers and learners continued doing what they normally do. Teachers in the intervention group were visited twice during the 14 weeks of implementation, when their compliance to the PMRP method was monitored. The same pre and post test was administered to both groups.

The original study found that schools which used the PMRP workbooks as planned made very significant gains compared to the control schools (Table 1). [2]

Table 1 Impact of learner performance with 11 weeks coverage*, Grade 6

	Pre-test	**Post-test**	**% Point Change**	***n (learners)***	**% Increase on BL**
Project Group	14.68	33.28	+18.60	1560	126.7%
Control Group	12.30	15.03	+ 3.0	1472	15.6%

Source: Schollar (2008)

* A number of teachers covered less than 80% of the 70 lessons contained in the material during the 14 weeks of the evaluation. These figures are for the learner scores of those teachers who had covered at least 11 weeks (78%) of the PMRP lessons.

The study concluded that the improvements in learner performance in intervention schools can be attributed to the PMRP approach. The gain scores achieved by the programme are in the order of twice to three times the kinds of learning gains effected by donor-funded school intervention programmes in South Africa in the last decade (Taylor, 2007).

4. Research Design and Methodology

Within the areas of school improvement in particular, JET Education Services has been working systematically over the past sixteen years to develop a comprehensive knowledge base on school effects and school improvement (see Taylor, op cit for a synthesis). Other initiatives in the country have focused on knowledge development through analysis of data produced by regional, national and cross-national studies of quality (van der Berg, 2008). In addition, there continues to be a large grouping of researchers undertaking qualitative research in the field. (For an overview of education research on learning achievement in South Africa, see Fleisch, 2008).

While these are important developments, what has been missing in South Africa is evidence-based research generated from randomized experiments. Although evidence generated from matched quasi-experimental studies (of which there have been a number) has the potential to make substantial contributions to the knowledge base, there are significant advantages to using randomized experiments, and as a consequence this is a rapidly growing area of research, particularly in the US (Raudensbush, 2005; Slavin, 2008).

The development of new statistical techniques now makes randomized controlled experiments both economically and practically feasible (EEPA, 2007). Consequently, there is a growing acceptance amongst education researchers that randomised experiments have an important role, particularly as they allow researchers to establish with a great degree of certainty the efficacy of specific programme interventions. The core value of these types of studies, if designed systematically and implemented with fidelity, is that they eliminate bias associated with uncontrolled factors.

Until very recently, such studies were uncommon in developing country contexts.   While the findings of these randomised experiments are clearly important, given the high-stakes consequences of their findings, it is necessary to expand the number of studies using these approaches and compare findings. One of the problems with some of the few existing studies is that the evaluations have often been undertaken by the programme developers, potentially compromising the independence of the investigations.

Randomised control trials have the following features:

- A relatively large sample of schools (around 40 or more).
- Randomly assigning schools in the sample to intervention and control groups, respectively. This is the key step, the bedrock assumption of the method, and ensures that all the many variables which influence learner performance are factored out of the equation. This step has earned the method the title of 'randomised control trial', which is the standard method in medical research and many other fields of enquiry. The method has not been employed to any great extent in the education field to date, although there is a rapidly growing literature on this approach in the US.
- Applying a 'treatment' to the intervention group (e.g. new kind of workbook), while the control group receives either the standard approach (or enhanced standard), or an alternative set of materials or method of instruction.
- Administering pre- and post-tests to both groups, and calculating the relative gain score of the intervention group.
- Any statistically significant gains registered by the intervention group can then confidently be attributed to the 'treatment'.

The PMRP Study used a modified standard treatment/control approach. For ethical and policy reasons, it was decided that rather than simply comparing the relative gains of intervention schools, compared with schools who received no intervention, the control schools would receive a complete set of materials and a limited amount of training, representing not what currently exists, but what could be standard practice if the schools was properly provisioned with existing approved materials and used these regularly. This is what we refer to as 'enhanced standard practice'.

Research site

Gauteng province was chosen as the research site. This has the advantages of containing a range of poor schools of different types (rural, urban, informal and formal) in relative proximity, which would reduce travel costs.

Sample design, frame and size

Particular care was taken in selecting the most appropriate sample design and sample size for the study.   The original study used a simple two group comparative design model, with twenty schools in the experimental group and twenty in the control. On closer examination, it was found that the decision made in the original study to place the number of schools at twenty for control and experimental groups respectively was not based on scientific sampling procedures, but rather on budget constraints.   Working closely with the University of Michigan's Capacity Building for Group

Level Interventions, the present study initially considered using a multisite cluster randomised field trial design (Borman, 2008). This is a novel approach, which essentially consists of working in relatively large schools and randomly assigning all classes *within* each school to the treatment or control, respectively, rather than using only one class per school for either treatment or control, as Schollar had done in his initial evaluation of the PMRP. The multisite cluster design has a number of advantages, chief of which is that fewer schools are needed to achieve the required number of classes, and hence this design holds the promise of reducing costs significantly. For example in Borman's (2008) evaluation of Open Court, the team used 17 classes spread over only 5 schools.

We proceeded to select schools according to the multisite cluster design, using all primary schools in Gauteng with four or more Grade 6 classes as the target population. However, when we began to approach schools, we realised that it is common for one teacher to be given responsibility for all Grade 6 maths classes in the school, thus nullifying the advantage of selecting more than one unique class in each school. We decided therefore to work in schools with a maximum of 2 Grade 6 classes, and adopted the whole school rather than classrooms within a school as the unit for random selection. The advantage of this design is that it reduces the possibility of spill-over or contamination between treatment and control classes, a serious problem in a pilot study of the Singapore Mathematics Materials in Alexander Township (du Toit, 2010).

For ethical and practical reasons, we could not sample intact classrooms within the treatment and control schools. In other words, all learners in a particular grade in a selected school were included in the study. The ethical reason is that sampling classrooms within schools would mean that some children would receive the benefits of the treatment or control within a single school and grade, others will not. The practical reason is that if the study had a sub-sample for treatment or control within a school, the mathematics teacher would have be required to teach two different methods simultaneously, which would substantially add to the workload.

With a random assignment to intervention schools and enhanced standard practice schools, the variance estimates are large because the schools were the unit of interest rather than the classroom (if classrooms were the unit of analysis, individual schools could have more than one unit). One of the vexing questions the researchers grappled with was the number of schools required to ensure that the study could have adequate statistical power.

Using Schochet's (2008) criteria, the study team made the following assumptions:

1. A single mathematics teacher for all Grade 6 classes in each school.
2. Only schools with two or fewer Grade 6 classes included.
3. Grade 6 mathematics taught through the medium of English.
4. Top wealthiest quintile of schools excluded.
5. 80% power level and 5% significance levels[3].
6. Testing restricted to a single grade.
7. ICC value (between school variance as a proportion of total variance) of 0.20[4].
8. Equal number of treatment and control schools.
9. All learners in a grade tested.
10. Attrition rates of less than 20%.
11. Minimum detectable effects (MDE) set at .33 of a standard deviation[5].
12. A survey of learners and schools to ensure R^2=.5.

Given these assumptions, according to Schochet, a sample size of 44 (22 treatment and 22 control) schools would be adequate, assuming a two-tail test, a value of .20 for the ICC, a balanced allocation of the research groups, and no subsample sample within units.

A list of all Gauteng public primary schools was drawn from the provincial database. All schools with more than two Grade 6 classes, fewer than 30 learners in Grade 6[6], and all schools in quintile 5 (i.e. most affluent schools) were excluded. The remaining schools were ordered randomly, and intervention and control schools assigned alternately off the random list. Starting at the top of this list, schools were then called and informed of the study. Where schools could not be contacted after repeated efforts, it was deleted from the list: of the first 53 schools on the list, 9 could not be contacted. In the main study in 2010, 21 intervention and 21 control schools had consent to participate.[7]

Pilot study

The main study was preceded by a pilot study covering 2 weeks in the selected schools in the third quarter of 2009. A defined area of work – fractions – was chosen and teachers trained to use the respective sets of materials. The purpose of the pilot was to test the logistical, pedagogical and analytical procedures to be used in the main study. Mean gains of 7,5% from pre- to post test in the PM group and of 6,2 % from pre- to post test in the CM group were both statistically significant.

Learner Attrition

Learner attrition of 11% occurred in both groups, because of learners not being present for one of the two tests, as shown in Table 2.

Table 2 Learner Attrition

CM*	Learners excluded	153
	Learners present both pre and post	1374
PM**	Learners excluded	125
	Learners present both pre and post	1146

Note: *CM (Classroom Mathematics) refers group that received the enhanced standard treatment.
**PM refers to the group that received the Primary Mathematics Research Programme materials.

The control group

In studies of this kind it is generally desirable to have a counterfactual: this is what the subjects of the experimental evaluation would have received if the intervention had not taken place. In the majority of South African schools 'normal schooling' consists of very slow pacing of instruction, and very little reading and writing, which is what Schollar (2008) used as a control group in his original study. However, it seems obvious that any group of classes which is subject to an intervention consisting of a set of materials which facilitates appropriate pacing and daily reading and writing would show improved learning, even if the materials were of a~~n~~ relatively mediocre quality. Therefore it was decided to provide control schools with additional standard materials, thus making it a comparison of the intervention with control schools using 'enhanced standard practice'.

Audit of Learner Materials

In the last week of May 2009, the study team undertook a comprehensive audit of all learner materials for the teaching of Grade 6 mathematics in the proposed control schools. The results are shown in Table 3. What emerged from the audit is that the randomly selected primary schools were using 11 different approved titles including *Successful Mathematics*, *Mathematics Plus*, *Spot On*, *Mathematics Matters, Classroom Mathematics, Maths for All, Maths Today, Magic Maths, My Clever Mathematics, Day by Day Mathematics, and Maths for Fun*. The most frequently used were *Successful Mathematics* (Oxford) and *Classroom Mathematics* (Heinemann).

Table 3 Audit Results of the Learner Support Materials in Control Group Schools

Name of school	Number of classes	Number of learners	Textbook(s) used	Number of books available	Shortage
S1	3	125	Successful Mathematics		
S2	1	72	Mathematics Plus	22	50
S3	2	64	Spot On Classroom Mathematics	1	64
S4	2	78	Mathematics Plus Mathematics Matters	3	They use photo copies (59)
S5	2	59	Classroom Mathematics	1	59
S6	1	12	Maths Today	12	0
S7	2	123	Maths for All	40	83
S8	2	78	Successful Mathematics		
S9	2	61	Successful Mathematics Spot On	Teachers copy	60
S10	1	48	Maths Today Magic Maths	5	43
S11	2	55	Successful Mathematics Spot On	0	55
S12	1	44	Spot On Understanding Maths	Teachers copy	44
S13	2	97	Classroom Mathematics Successful Mathematics	20	77
S14	2	83	Successful Mathematics	0	83
S15	1	51	Successful Mathematics Spot On	51	0
S16	1	51	My Clever Mathematics Classroom Mathematics	Teachers copy	51
S17	1	46	Spot On	Teachers copy	46
S18	1	41	Classroom Mathematics	35	6
S19	2	75	My Clever Mathematics	38	37
S20	2	88	Classroom Mathematics Day by Day Mathematics My Clever Mathematics	7	81
S21	2	90	Magic Maths Spot On	Teachers copy	90
S22	3	120	Spot On Maths for fun	0	120

The other major finding from the audit was that all but two of the schools had textbook shortages. Half of the schools had only a teacher's copy of the book. The learners in these schools did not have access to a single copy of the book. In the rest of the schools, most learners had to share textbooks.

Choice of materials for the Control group

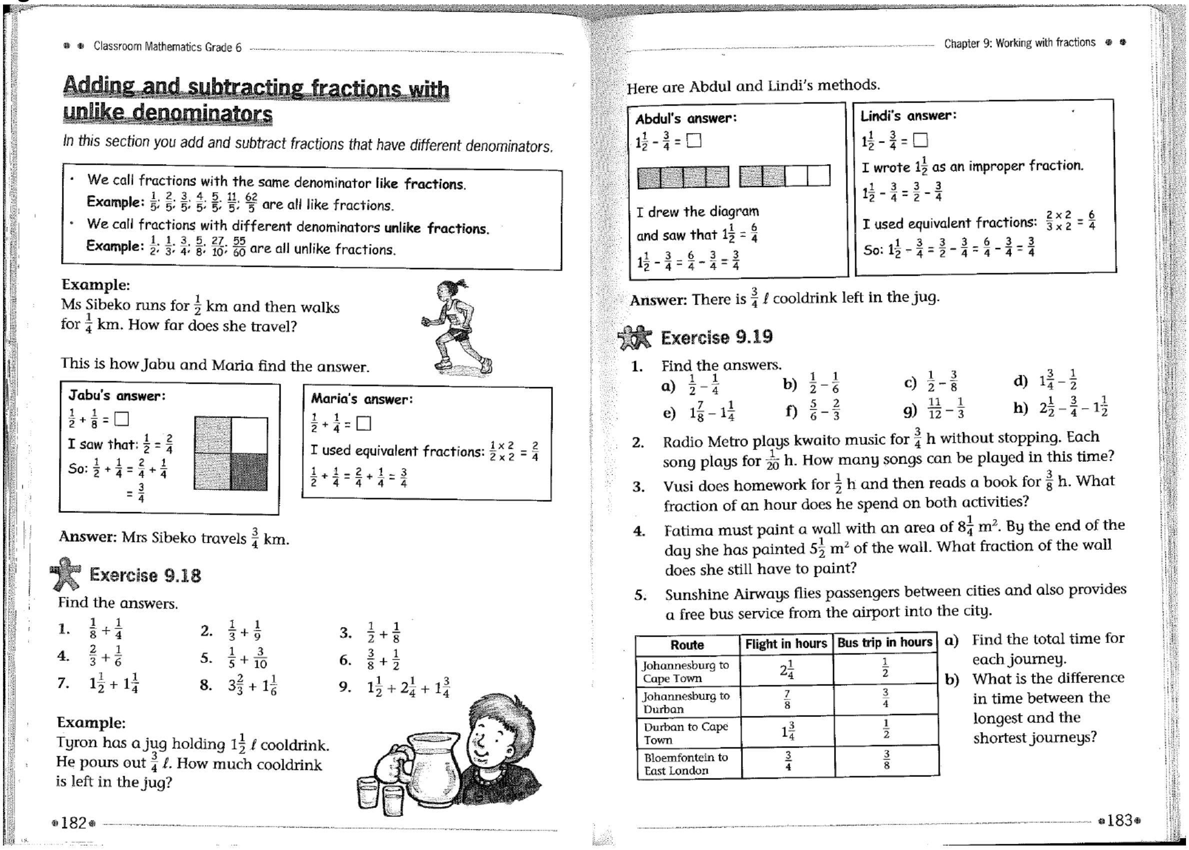
The results of the materials audit in control schools revealed that the schools did not have sufficient materials for the study of mathematics in the Intermediate Phase (and by implication for all levels of the primary school). Thus, implementation of the research design described above (PMRP vs 'enhanced standard practice') would not be possible without the provision of books to all learners in control schools. The *Classroom Mathematics* textbook was chosen for the control group, on the grounds that it is widely regarded as a 'good' textbook and is very widely used throughout the country. A comprehensive set of books were provided to schools during training in the second week of 2010.

The two sets of materials used by the teachers and their learners are designed to cover LO1 of the NCS, but in different ways. The following review gives some indication of the similarities and differences between the material sets.

Classroom Mathematics (CM)

The book used by learners in this group was *Classroom Mathematics: Grade 6* (Scheiber, et al, 2009). The CM material set is a standard textbook approved for use in South African schools. It was developed by an experienced mathematics textbook writing team of seven writers.

Figure 1: A lesson from Classroom Mathematics



Classroom Mathematics Grade 6

Adding and subtracting fractions with unlike denominators

In this section you add and subtract fractions that have different denominators.

- We call fractions with the same denominator **like fractions**.
  Example: $\frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}, \frac{5}{5}, \frac{11}{5}, \frac{62}{5}$ are all like fractions.
- We call fractions with different denominators **unlike fractions**.
  Example: $\frac{1}{2}, \frac{1}{3}, \frac{3}{4}, \frac{5}{8}, \frac{27}{10}, \frac{55}{60}$ are all unlike fractions.

Example:
Ms Sibeko runs for $\frac{1}{2}$ km and then walks for $\frac{1}{4}$ km. How far does she travel?

This is how Jabu and Maria find the answer.

Jabu's answer:
$\frac{1}{2} + \frac{1}{8} = \square$
I saw that: $\frac{1}{2} = \frac{2}{4}$
So: $\frac{1}{2} + \frac{1}{4} = \frac{2}{4} + \frac{1}{4} = \frac{3}{4}$

Maria's answer:
$\frac{1}{2} + \frac{1}{4} = \square$
I used equivalent fractions: $\frac{1 \times 2}{2 \times 2} = \frac{2}{4}$
$\frac{1}{2} + \frac{1}{4} = \frac{2}{4} + \frac{1}{4} = \frac{3}{4}$

Answer: Mrs Sibeko travels $\frac{3}{4}$ km.

Exercise 9.18

Find the answers.

1. $\frac{1}{8} + \frac{1}{4}$
2. $\frac{1}{3} + \frac{1}{9}$
3. $\frac{1}{2} + \frac{1}{8}$
4. $\frac{2}{3} + \frac{1}{6}$
5. $\frac{1}{5} + \frac{3}{10}$
6. $\frac{3}{8} + \frac{1}{2}$
7. $1\frac{1}{2} + 1\frac{1}{4}$
8. $3\frac{2}{3} + 1\frac{1}{6}$
9. $1\frac{1}{2} + 2\frac{1}{4} + 1\frac{3}{4}$

Example:
Tyron has a jug holding $1\frac{1}{2}$ ℓ cooldrink. He pours out $\frac{3}{4}$ ℓ. How much cooldrink is left in the jug?

182

Chapter 9: Working with fractions

Here are Abdul and Lindi's methods.

Abdul's answer:
$1\frac{1}{2} - \frac{3}{4} = \square$
I drew the diagram and saw that $1\frac{1}{2} = \frac{6}{4}$
$1\frac{1}{2} - \frac{3}{4} = \frac{6}{4} - \frac{3}{4} = \frac{3}{4}$

Lindi's answer:
$1\frac{1}{2} - \frac{3}{4} = \square$
I wrote $1\frac{1}{2}$ as an improper fraction.
$1\frac{1}{2} - \frac{3}{4} = \frac{3}{2} - \frac{3}{4}$
I used equivalent fractions: $\frac{2 \times 2}{3 \times 2} = \frac{6}{4}$
So: $1\frac{1}{2} - \frac{3}{4} = \frac{3}{2} - \frac{3}{4} = \frac{6}{4} - \frac{3}{4} = \frac{3}{4}$

Answer: There is $\frac{3}{4}$ ℓ cooldrink left in the jug.

Exercise 9.19

1. Find the answers.
   a) $\frac{1}{2} - \frac{1}{4}$ b) $\frac{1}{2} - \frac{1}{6}$ c) $\frac{1}{2} - \frac{3}{8}$ d) $1\frac{3}{4} - \frac{1}{2}$
   e) $1\frac{7}{8} - 1\frac{1}{4}$ f) $\frac{5}{6} - \frac{2}{3}$ g) $\frac{11}{12} - \frac{1}{3}$ h) $2\frac{1}{2} - \frac{3}{4} - 1\frac{1}{2}$
2. Radio Metro plays kwaito music for $\frac{3}{4}$ h without stopping. Each song plays for $\frac{1}{20}$ h. How many songs can be played in this time?
3. Vusi does homework for $\frac{1}{2}$ h and then reads a book for $\frac{3}{8}$ h. What fraction of an hour does he spend on both activities?
4. Fatima must paint a wall with an area of $8\frac{1}{4}$ m^2. By the end of the day she has painted $5\frac{1}{2}$ m^2 of the wall. What fraction of the wall does she still have to paint?
5. Sunshine Airways flies passengers between cities and also provides a free bus service from the airport into the city.

Route	Flight in hours	Bus trip in hours
Johannesburg to Cape Town	$2\frac{1}{4}$	$\frac{1}{2}$
Johannesburg to Durban	$\frac{7}{8}$	$\frac{3}{4}$
Durban to Cape Town	$1\frac{3}{4}$	$\frac{1}{2}$
Bloemfontein to East London	$\frac{3}{4}$	$\frac{3}{8}$

a) Find the total time for each journey.
b) What is the difference in time between the longest and the shortest journeys?

183

CM is generally divided into two-page 'chunks' of work, each of which is structured as follows: introduction of terminology, explanation of a concept or technique, a number of worked examples, and a set of activities (generally in the form of an exercise) which are graded from easy to more difficult. The activities are designed to be done in separate exercise books. Although the teacher guide contains supplementary material and solutions to all problems, the book is self-sufficient and, therefore suitable for self-study by learners. A typical two-page spread from CM is shown in Figure 1.

CM provides thorough curriculum coverage with many worked examples, good theoretical definitions and a wide spread of exercises for learners ranging from repetitive numeric activities to contextualised word problems. Both the Learners' Book and the Educator's Guide are written in accessible language. In the Educator's Guide the teacher is given detailed explanations of the current curriculum and ways in which the Learners' Book will enable her to meet the teaching requirements of the national curriculum. It also contains full worked solutions to all activities in the learner books; however, in general, guidance does not follow the Learner Book in detail, being of a more general nature about the topics under discussion.

The Learners' Book is designed to cover the entire NCS mathematics curriculum in a format that can be used by a teacher over the period of one year. A work schedule which could be followed by teachers is provided but no emphasis is placed on the use of this work schedule. The Learners' Book also contains expositions of alternate algorithms, such as using expanded notation to 'decompress' 'long' multiplication and division: some of these examples are long and unwieldy, and while they can be an aid to conceptual understanding, it would be inefficient to use them to perform operations, particularly for large numbers. CM strongly promotes problem solving and does encourage the drilling of multiplication tables and addition/subtraction bonds.

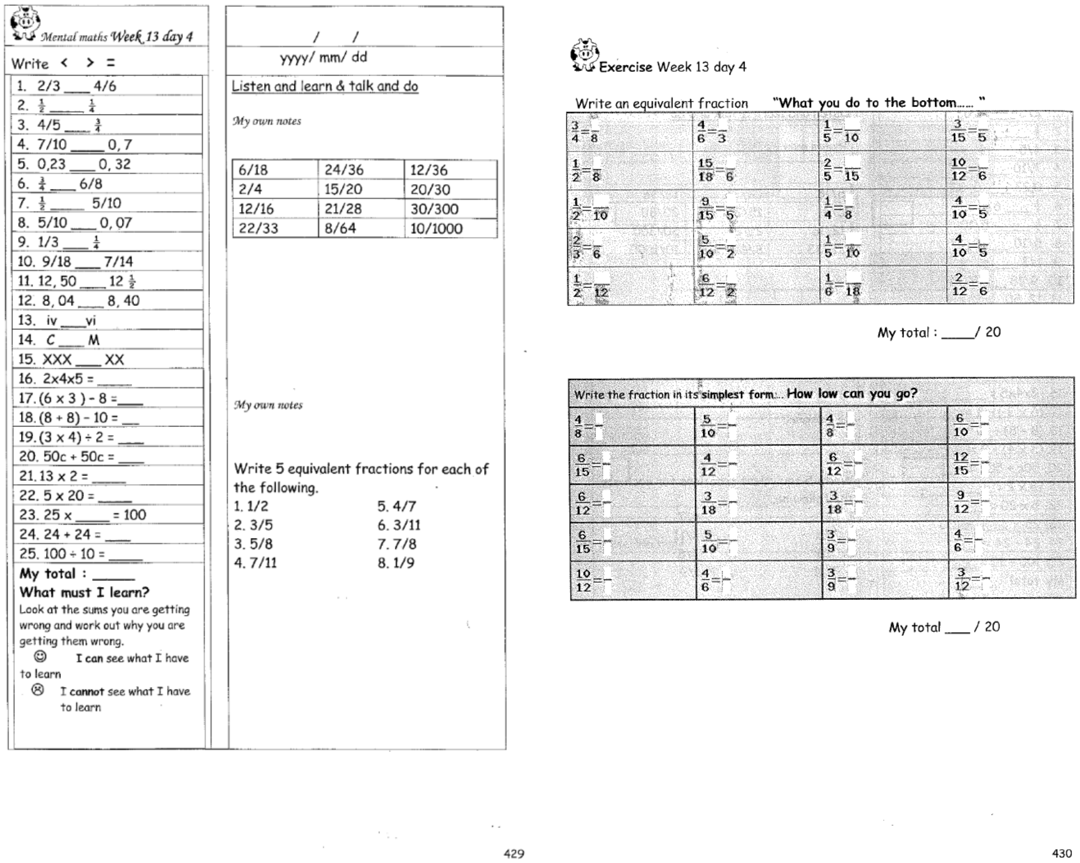
Primary Mathematics Research Project (PM)

The book used by learners in this group was *Back to Basics! Getting Learning Outcome One Right Intermediate Phase* (Schollar, 2003). The PM material set is a project learner workbook and teacher manual that has been used in several South African schools. The book was explicitly designed to address the problems observed during extensive research in primary schools, specifically to move beyond the 'unit counting' methods so ubiquitous in schools serving poor children and to develop a greater degree of automaticity in executing the four arithmetic operations. PM's approach is not to prioritise basic operations at the exclusion of problem solving, but assumes that proficiency in the four operations is a prerequisite for problem solving. In order to achieve these aims, the workbooks consist of exercises which are mostly repetitive, numeric activities designed to consolidate concepts that have been taught. Each day starts with a 10 minute mental maths exercise. Contextualised word problems are also provided at the end of the fourth and the ninth weeks of the materials and the formal assessment tasks include some word problems. The design of PM is illustrated in the two-page extract shown in Figure 2.

PM provides fairly thorough curriculum coverage of LO1 (it is not a full "textbook" that covers the entire curriculum). The materials require that teachers test their learners and then place them into strands according to their achievement on the diagnostic test. Each level provides the learner with the same mathematical content but on a different level ranging from Grade 3 level work to Grade 6 level work. This feature of the book may make it complex to use, and certainly increases its size: this is a very bulky book.

Both the Learner Workbook and the Teacher Manual are written in accessible language. In the Teacher Manual a discussion of the current crisis in mathematics education in schools and some possible explanations for this crisis are given. The PM materials approach is explained in detail to the teacher, and the work schedule to be followed by teachers is provided. The manual also gives input on the value of mental maths and the importance of routine algorithms. The guide challenges the NCS requirement for the use of alternative algorithms, suggesting that the strong emphasis on these algorithms, some of which are very cumbersome to use, can confuse children and may even encourage unit counting and other inefficient methods. The teacher manual calls on teachers to commit to teaching using the PM materials diligently, according to the given work schedule, with the assurance that their learners' mathematics will be improved.

Figure 2: Primary Mathematics Research Programme learner materials



Mental maths Week 13 day 4

Write < > =

1. 2/3 ___ 4/6
2. ½ ____ ¼
3. 4/5 ___ ¾
4. 7/10 ____ 0,7
5. 0,23 ___ 0, 32
6. ¾ ___ 6/8
7. ½ ____ 5/10
8. 5/10 ___ 0, 07
9. 1/3 ___ ¼
10. 9/18 ___ 7/14
11. 12, 50 ___ 12 ½
12. 8, 04 ___ 8, 40
13. iv ___ vi
14. C ___ M
15. XXX ___ XX
16. 2x4x5 = _____
17. (6 x 3) - 8 = ____
18. (8 + 8) - 10 = ___
19. (3 x 4) ÷ 2 = ____
20. 50c + 50c = ____
21. 13 x 2 = _____
22. 5 x 20 = _____
23. 25 x _____ = 100
24. 24 + 24 = ____
25. 100 ÷ 10 = _____

My total : _____

What must I learn?

Look at the sums you are getting wrong and work out why you are getting them wrong.

☺ I can see what I have to learn

☹ I **cannot** see what I have to learn

/ /

yyyy/ mm/ dd

Listen and learn & talk and do

My own notes

6/18	24/36	12/36
2/4	15/20	20/30
12/16	21/28	30/300
22/33	8/64	10/1000

My own notes

Write 5 equivalent fractions for each of the following.

1. 1/2 5. 4/7
2. 3/5 6. 3/11
3. 5/8 7. 7/8
4. 7/11 8. 1/9

429

Exercise Week 13 day 4

Write an equivalent fraction **"What you do to the bottom….. "**

3/4 = /8	4/6 = /3	1/5 = /10	3/15 = /5
1/2 = /8	15/18 = /6	2/5 = /15	10/12 = /6
1/2 = /10	9/15 = /5	1/4 = /8	4/10 = /5
2/3 = /6	5/10 = /2	1/5 = /10	4/10 = /5
1/2 = /12	6/12 = /2	1/6 = /18	2/12 = /6

My total : _____/ 20

Write the fraction in its simplest form… **How low can you go?**			
4/8 =	5/10 =	4/8 =	6/10 =
6/15 =	4/12 =	6/12 =	12/15 =
6/12 =	3/18 =	3/18 =	9/12 =
6/15 =	5/10 =	3/9 =	4/6 =
10/12 =	4/6 =	3/9 =	3/12 =

My total ____ / 20

430

The teacher guide and learner book are designed to be used in tandem, and the learner books are therefore not amenable to self-study, containing mainly mental arithmetic exercises, rule summaries and sets of written exercises designed to be done directly in the workbook. The teacher guide contains full lesson plans, with conceptual explanations, definitions and worked examples, and follows the learner book closely, even containing advice to the teacher on what to write on the board during each lesson.

Both sets of materials are very 'busy', characterised by crowded pages and cannot be easy for Grade 6 readers. CM is very 'wordy', although the text is somewhat leavened by the inclusion of drawings and cartoon figures who talk to the reader through speech bubbles. PM contains fewer words and explanations, but is crowded with exercises.

Test construction

A test covering the chosen LO for the grade was constructed by the research team. The first draft of the test consisted of 60 items covering eight skill categories – number concepts (place value, comparing numbers), fractions, addition, subtraction, multiplication, division, problem-solving and mental operations – all specified in the NCS for Grade -3 to 6. This test was administered to both project and control groups as a pre- and post-test. Item difficulty and discrimination indices were calculated per group for each item at the pre- and post-test. Using the results of the post-test, 40 items were selected with item means between 0.15 and 0.91 and item discrimination indices between 0.20 and 0.80 on the post-test.

Furthermore, the reliability estimates for the whole sample as well as each group were calculated on both the pre-and post-test. The measure of reliability gives an indication of the likelihood of obtaining the same results should the test be administered again under similar conditions. It is usually measured by the Alpha coefficient. The overall reliability of the test, as indicated by the Alpha coefficient, was 0.84 (CM 0.84 and PM 0.85) at the pre-test and 0.87 (CM 0.87 and PM 0.87) at the post-test. By convention, a lenient cut-off of .60 is common in exploratory research; Alpha should be at least .70 or higher to retain an item in an "adequate" scale, and many researchers require a cut-off of .80 for a "good scale" (Gronlund, 1998). Therefore, with an upper level of 1, the coefficient of 0.84 and 0.87 is high and would suggest that the test had good reliability at both the pre-and post-test level.

Table 4 Reliability Statistics for the Pre and Post Test

Cronbach's Alpha		
Group	**Pre-test**	**Post-test**
CM	.842	.871
PM	.845	.870

The distribution of the items across skill categories is shown in Table 5.

The 40 items that remained in the test for analysis were spread across the eight skills categories and Grades 3, 4, 5 and 6 as indicated above. Grade level does not necessarily indicate difficulty level, since, for example, a Grade 6 item might be pitched at the minimum requirement level for that grade, while a Grade 5 item might be at the maximum level. The skills categories can be further grouped into operational concepts and procedures (addition, subtraction, multiplication and division – these questions were predominantly procedural), number concept and understanding (number and fractions – these questions were predominantly conceptual) and strategic (problem solving – these questions tested the ability of learners to interpret questions that applied their number or operational skills in context).

Table 5 Test items in grouped according to Skill and Grade level*

Number of items					
Skill	Grade 3	Grade 4	Grade 5	Grade 6	TOTAL
Number concepts	3	3	1	1	8
Fractions	3	3	1	1	8
Problem solving	4	2	0	0	6
Addition	1	2	0	1	4
Subtraction	1	1	2	0	4
Multiplication	3	0	0	0	3
Division	2	1	0	1	4
Mental Operations	1	1	0	1	3
TOTAL	18	13	4	5	40

* Grade levels follow the NCS specification and do not necessarily indicate the difficulty of an item.

Items can be categorised as those containing only symbolic information (numbers and mathematical signs) (10), as questions involving a reading instruction with numeric information (24), or as problem type questions with the numeric information embedded in the text (6). Achievement on these items could be compared to investigate the contention that learners struggle with reading and problem solving activities.

The test items were scored according to a key which included some items for which there was only one correct answer and some items that were partially scored. The partial scoring gave recognition to the correct answer (of which there was one) but also to answers which indicated a partial understanding of or solution to the question. Tracking the improvement from incorrect, to partially and fully correct could give a deeper understanding of learner gains over the 14 week teaching period. There were 33 items with one correct answer only and 7 items with partial scoring.

Training of Teachers

Training took place over two weekends in January. One of the researchers, who is an experienced teacher educator, trained teachers for both the PM and the CM groups. This was planned so that teachers had a few days to complete the necessary administrative tasks required of them at the beginning of the academic year, before beginning teaching according to the 14 week schedule. It was also hoped that by the second week of term most of the learners would have registered at their schools for the year and numbers of learners in each class would be known, since a key aspect of this project is the provision of a textbook to each learner in every class. Both groups received the same number of hours of training, although the content of their training differed according to differences in the materials being used by the group. Teachers were told that attendance at the training sessions was compulsory, but there were still certain teachers who were unable to attend due to various reasons. These teachers were visited at their schools during the week between the first and the second training sessions. This visit was to deliver essential materials to the teachers that had been handed out at the first session and to brief them on what they had missed in the first session. Teachers who were unable to attend the second training session were also visited by a fieldworker and their learner materials were delivered to them.

5. Findings

The findings section is divided into four sections. The first section, responding to the literature, explores the gains made by learners between the pre- and post-tests. Although the study was not explicitly designed to contribute to the debate on the relative effectiveness of providing materials *per se* (as no conventional control group was used), the evidence does provide indirect support for the value of the provision and systematic use of learner materials. The second section addresses the core concern of the study, the relative effectiveness of the Primary Mathematics Research Project approach, and specifically the universal provision and use of the workbook compared to an off-the-shelf mathematics textbook. To this end, we compared pre- and post-test scores, percentage point gains, percentage gains of the baseline score, and gains in relation to curriculum coverage. The third section narrows the focus by exploring relationships between the respective materials' approaches to mathematics and gains on the test scores. Here we investigate both the relative success of materials in improving specific mathematics skills, and the extent to which these skill improvements are conditioned by the relative academic strength of groups of learners on the pre-test. Finally, while the study had not intended to identify specific challenges involved in the teaching of intermediate phase mathematics, error analysis of a number of items suggests interesting trends that could be taken up by researchers in the field. The last section is of a preliminary nature and would benefit from a more extensive analysis, guided by the voluminous literature on this topic.

5.1. Gains in the sample population associated with the provision of learner materials

This section responds to the wider debate about textbooks or learner resources and their contribution to improved learning outcomes. The two interventions, while providing different kinds of learning materials, share a common emphasis on universal provision and consistent use. At the most basic level, the results of the study show substantial gains in learner test scores in both the experimental and control groups. The gains are large by most conventional indices, including percentage gains in the average scores (Table 6).

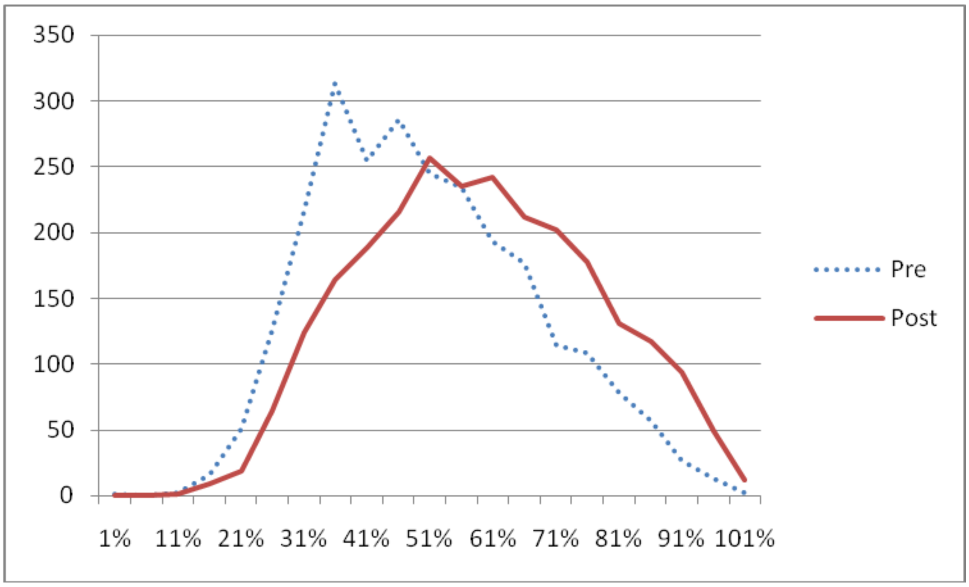
Table 6 Mean Scores and Gains for the Study Population, mean percent correct

	n	Minimum	Maximum	Mean	Std. Deviation
Pre-test	2515	.00	100.00	47.47	17.00
Post-test	2515	10.00	100.00	55.87	18.29
Gain	2515			8.40	11.54

The population (both groups combined) achieved a percentage point increase of 8.40%, which represents an improvement of 17.7% on the pre-test mean score. The learner materials (with minimal training and monitoring) are the most likely factor that could explain these gains. But as this was not the primary concern of this study, robust causal claims cannot be made. That said, average gains of above 5 percentage points for any education intervention is unusual both in the international and South African literature (Taylor, 2007).

Figure 3 shows the distribution of population scores on the pre- and post-tests. The distribution indicates the extent to which the gains occurred across the entire population.

Figure 3. Distribution of Learner Scores for Pre and Post Tests



One of the most striking statistics is the change in the proportion of learners correctly answering 50% or more of the test items. Of 2515 learners in the sample, only 1062 achieved 50% or above on the pre-test. By the post test, the number had increased to 1533, an improvement of 471 learners.

In other words, 18.7% more children in the population achieved a score of 50% or more in the post-test.

5.2. Relative gains for Primary Mathematics Research Project and Classroom Mathematics.

The study compares the PMRP materials with training and monitoring to 'enhanced standard practice', i.e. the use of a textbook for all learners with training and monitoring. We consider this to be 'enhanced' standard practice, since it is what is expected to happen in schools, but which has been shown by numerous research studies in South Africa to be very rare, with few books in existence in the large majority of classes, and little reading and writing featuring among learner activities. We selected the materials for the enhanced standard practice on the basis of what is widely in use and is regarded by subject experts as being of good quality.

Despite the differences in both form (workbook vs textbook) and internal logic (see Material Review above) there was no significant difference in the gains made by the two groups, each group achieving a mean improvement of over 8 percentage points (Table 7).

Table 7 Mean Scores in the Pre and Post Tests

	Number of learners		Mean for test				Std. Deviation	
			Mean % correct		Std. Error			
	CM	PM	CM	PM	CM	PM	CM	PM
Pre-test	1374	1141	48.07*	46.74*	.45	.50	16.89	17.091
Post-test	1374	1141	56.69*	54.87*	.49	.54	18.25	18.29
Gain			8.62**	8.13**				

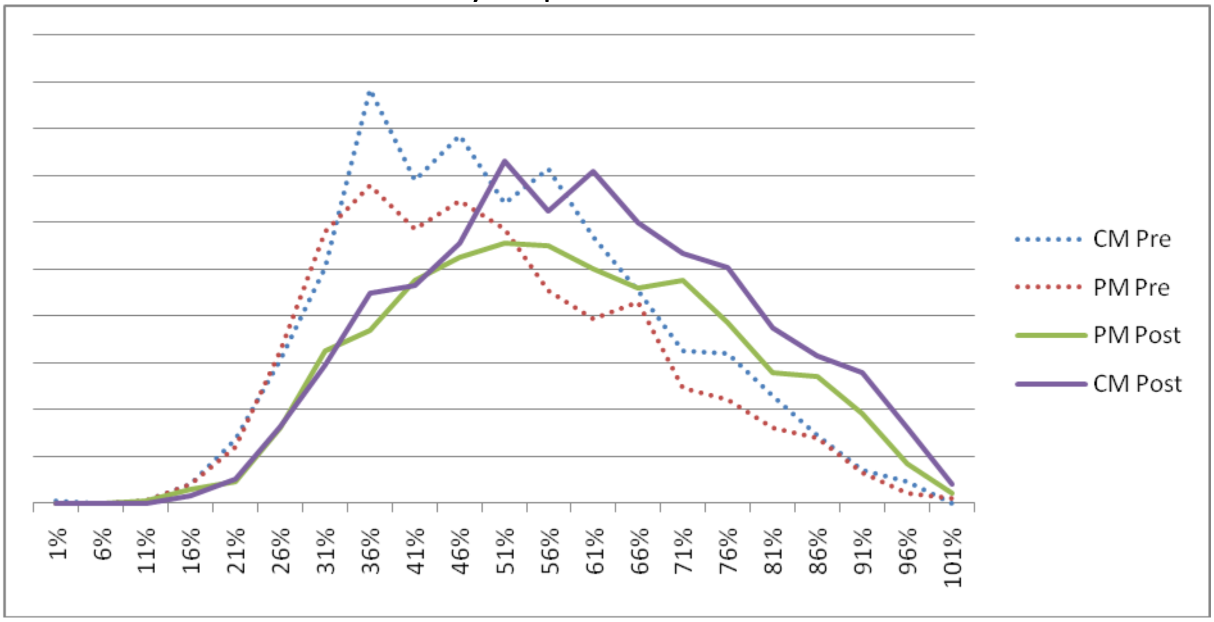
* The difference between the pre- and post-test mean for each group is statistically significant, as is the gain score within each group (p-value of .00)

* * The difference in gains between CM and PM is not statistically significant (p-value of 0.29).

Although the schools were selected at random, the PM group performed less well on the pre-test. However the percentage point gains were almost identical.

The respective frequency distributions of the two groups in both pre- and post-test are shown in Figure 4.

Figure 4 Pre- and Post-test score distribution by Group



Learners using both sets of materials made substantial gains in terms of proportions achieving scores of 50% or more (Table 8). Nearly one-fifth more children in both groups achieved this 'pass' on the post-test, which constitutes nearly 45% more than had passed the pre-test. Whether assessed on the percentage point gains, on the percentage gains from the pre-test baseline or the increased proportion of learners that 'passed' the test, there is little evidence that either set of materials is more advantageous. Both made significant gains of roughly the same magnitude.

Table 8 Proportion of learners scoring at or above 50%

Group	Number of learners	Pre-test		Post-test		Increase		
		No. ≥50%	Perc ≥50%	No. ≥50%	Perc ≥50%	No. ≥50%	Perc ≥50%	Perc on BL
CM	1374	604	43.96	866	63.03	262	19.07	43.4%
PM	1141	458	40.14	667	58.46	209	18.32	45.6%

Slavin and Lake (2008) best evidence synthesis of elementary mathematics programmes revealed while the introduction of new textbooks alone did not seem to impact positively on learner outcomes, what they refer to as 'instructional process strategies' do. We would suggest that context is key as most disadvantaged South African schools do not have complete sets of textbooks or workbooks and as such, the provision and use of such learner materials is likely, at least in part, to explain the change in prevailing instructional practices.

Gain scores by coverage of the materials

Both sets of materials are designed to cover all LO1 topics specified for Grade 6 in the NCS in 14 weeks. The time allocated is proportionate to the space LO1 occupies in the curriculum. The study tracked coverage of the materials through to visits to classroom on three separate occasions.

Fieldworkers examined learner books and noted how far the class had progressed and whether all exercises had been completed up to that point. From this data, percentage coverage could be calculated for each class. The two interventions diverge in terms of the gain scores by degree of coverage (Table 9).

Table 9 Learner Mean Totals and Gains per Coverage at 79% Level

Coverage		Number		Mean score (%)		Std. Deviation	
		CM	PM	CM	PM	CM	PM
≥79%	Pre test	607	420	51.29	44.07	17.98	16.50
	Post test	607	420	63.29	52.91	17.57	19.01
	Gain	607	420	12.00	8.85	12.94	10.56
<79%	Percentage on Pre test	622	721	46.97	48.30	15.71	17.25
	Percentage on Post test	622	721	53.58	56.02	17.05	17.78
	Gain	622	721	6.61	7.72	11.61	10.16
Missing*	Percentage on Pre test	145		39.35		13.10	
	Percentage on Post test	145		42.45		13.92	
	Gain	145		3.09		9.87	

* Note: School was missed two out of the three visits.

The first point to note is that, whereas 607 (44%) learners in the CM group achieved 79% coverage or more, only 420 (37%) of the PM group had. This may be because CM is more easily readable than PM, and therefore classes are able to progress faster, although this hypothesis will have to be investigated by means of case studies. Also, those CM learners who covered 79% or more achieved higher gain scores (12 percentage points) than their PM counterparts (8.85). On the other hand, those PM learners who covered less that 79% of the material showed higher gains (7.72) than the equivalent fraction of CM learners (6.61). These differences are striking, but their origin is unclear and a great deal more qualitative work would be required to gain a better understanding of this phenomenon. In general, for both groups, gain scores increase with increasing coverage, an obvious and expected development, which provides strong evidence in support of the hypothesis that it is use of the materials that is causing learning gains.

Costs

To determine the unit costs of the respective materials, the study assumed the costs for the *Classroom Mathematics* and Primary Mathematics materials as paid to the publisher in January 2010.

Table 10 Unit Cost of Materials (R)

	CM	PM
Teacher Guide	145	136
Learner Book	87	83
Teacher Guide Cost Per Learner*	3.40	3.63
Average Cost of Learner Book per Year**	43.50	166
Total Annual Unit Costs	**R 46.90**	**R 169.63**

* Per Learner Costs of Teachers Guide: we divided the full costs of the teachers guide by 40.
** To establish the average cost of the learner book per year, we assumed that the CM book would be used for at least two years (this is a very conservative estimate that incorporates substantial loss). Primary Mathematics materials are designed to cover only half of the years work, and are designed to be replaced every year. We therefore multiplied the cost by two.

Given that Classroom Mathematics is a textbook, we assumed multi-year use, but given the difficulties with retrieval systems, we assumed that the books would need to be replaced every two years. With regard to the annual costing of the Primary Mathematics book, we assumed that this is a workbook in which learners would be writing directly, and as such they would need to be replaced annually. Moreover, the book is designed to cover only half of the NCS curriculum for Grade 6 mathematics, and as such, we assumed that a full year of this kind of materials would cost twice what we paid per book. Within the constraints of our assumptions, which we consider to be realistic, the PM materials are in the order of three times as costly as the CM textbook and accompanying teacher guide.

5.3. Improvement by Skill and Prior Academic Achievement

A close analysis of the data by mathematical skill and learner pre-test scores reveals important differences in achievement of the two groups. We divided learners in both groups into three achievement categories of roughly equal size: those scoring 0-37% on the pre-test, those achieving 38-54%, and the top learners who scored 55% or more. Comparisons of this kind are more valid for the top and bottom group because the division lines are arbitrary.

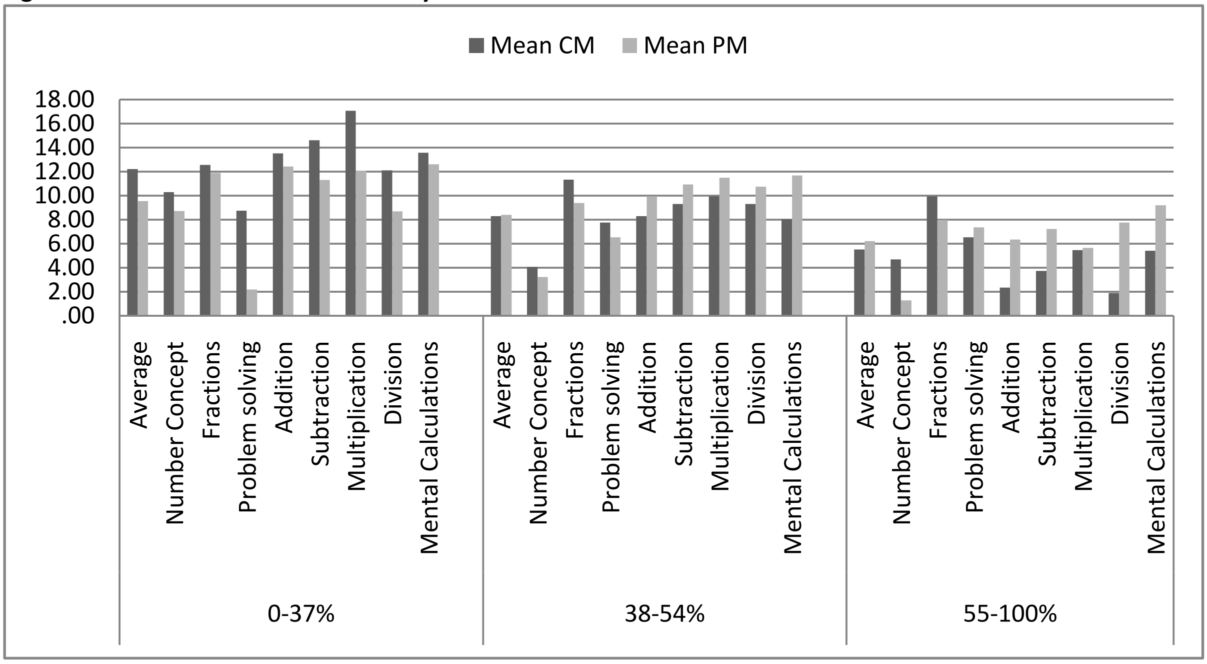
Table 11 Gains by Learner Category

Learner category	Test	*n*		Mean		Std. Deviation	
		CM	PM	CM	PM	CM	PM
0-37%	Pre-test	453	411	30.07	29.46	6.07	5.75
	Post-test	453	411	42.28	39.01	13.15	11.10
	Gain	453	411	12.21	9.55	13.39	10.55
38-54%	Pre-test	446	377	46.10	45.93	4.58	4.57
	Post-test	446	377	54.39	54.34	12.71	11.69
	Gain	446	377	8.28	8.41	12.07	10.70
55-100%	Pre-test	475	353	67.11	67.72	9.79	9.72
	Post-test	475	353	72.62	73.92	13.81	11.69
	Gain	475	353	5.52	6.20	10.95	9.31

The first point to note is that the CM group made stronger gains than the PM group among the weakest students, almost identical gains for the average group, and marginally lower gains for the strongest students (Table 11). This appears to be counter-intuitive as the PM materials were explicitly designed with weaker learners in mind. Here too, the reason for these differences is unclear.

A further disaggregation of learner gain scores by prior achievement and mathematical skill is illuminating (Figures 5). First we look at the lowest scoring group of learners, where Table 11 shows that CM learners outperformed their PM counterparts. Figure 5 shows that the CM advantage is spread across all skill categories.

Figure 5 Gains on Skills for Learners by Achievement on the Pretest



The pattern for learners who achieve in the middle range may be associated with the intentions of the two respective sets of materials designers. Figure 5 shows that PM learners outperformed CM learners in the four operations and in mental calculations, which is in accord with the design of PM. For the higher level skills – problem solving and fractions – CM shows an advantage, which is also congruent with the design intentions. This is not surprising given the focus on mental arithmetic (habituation) on these operations required by the PM materials: all lessons begin with mental mathematics warm-up exercises which focus almost exclusively on the four arithmetic operations.

Figure 5 shows the gains by skill for the top performing learners. These results are more mixed, but distinct residues of the patterns shown by lower achieving learners, with PM learners generally scoring better on the four operations and mental maths, and CM learners having an advantage in fractions and number concepts. However, a major difference between the top- and middle-achievers is that top PM students seemed to benefit more in the area of problem solving than CM learners. Perhaps the PM philosophy that problem solving becomes easier when the number bonds and multiples have been habituated is being vindicated. This is another tentative hypothesis, however, that would need to be investigated in much greater detail.

As noted earlier, the PMRP approach was developed as a response to the recognition that many children were using inappropriate concrete unit counting procedures to solve basic operation questions. A careful analysis of the results of the study show that in terms of this specific aspect of the mathematics, clearly a key component, the materials is effective. Furthermore, the data reveals that the PMRP approach is most effective at enhancing basic operational skills for children with average to higher than average ability.

5.4. Error Analysis

There is an auxiliary value to the PMRP's contribution. In addition to identifying possible patterns in terms of which skills are best taught by which materials/methods, the study also makes an important contribution to error (misconceptions) analysis. This builds on and complements the work of initiatives such as the Data Informed Practice Improvement Project (DIPIP). While Schollar's original insight focused on the relative prevalence of unit counting and the difficulties that teachers and learners had with transitioning to abstract reasoning and the systematic use of standard algorithms, the specific questions support quantitative confirmation of a number of mathematics insights.

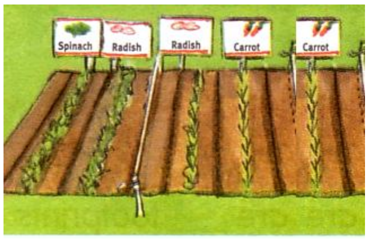
For example, the importance of language and reading is highlighted in the difference between questions 27 and 32. In the post-test more than 60% of learners selected the correct answer to item 27 but only 30% of learners selected the correct answer to item 32.

(27.) 20 x 5 = ____

(32.) Emma makes a garden. She plants 20 seeds in a row.

She makes 5 rows.

How many seeds does she plant altogether?



Comparative analysis of other questions reveals other common errors associated with learner unfamiliarity with the meaning of the arithmetic symbols. +, -, x, and ÷. In question 11, approximately 10% of learners selected 80 as the correct answer. These learners probably used addition instead of subtraction.

(11.) 68 – 12 = ______

Specific questions reveal a great deal about the problem of over-generalisation. See for example, the problem that children had with item 6.

(6.) $\frac{1}{2} + \frac{1}{2} =$ ____

The most popular distracter selected as the correct solution to this problem was **C**. $\frac{2}{4}$. Error analysis here reveals that learners were applying an over-generalisation of the "rule" – "what you do to the top you do to the bottom". This has implications for materials developers.

These three examples illustrate the value of error analysis, and learner responses to all questions in the test would benefit from further study.

6. Conclusion and implications of the study

Research

The advantage of random control trial (RCT) designs, as used in the present study, is that they are able to generalize research findings to the target population, in this case Grade 6 classes in quintile 1-4 schools in Gauteng. And the answer to our main research question – does the PMRP, as a customised workbook – show an advantage over a conventional textbook – is clearly in the negative.

However, the conclusions that can be drawn from RCTs are constrained by the elements of their design. In the case of our study, the absence of a control group – for both ethical and practical reasons – means that we cannot compare the results of the present study with what normally happens in Gauteng schools. We can show relatively large learning gains, but we do not know how much of this might have happened without the materials provided by the research study. However, we can make an educated guess, based on typical gains made by school development programmes over the last 20 years, which indicate that all learners in our study gained significant advantage through using the materials.

A second constraint on RCT research designs is that they are blind to the mechanisms which drive the changes observed. In other words, they can confidently show a change accompanying an intervention, but, on their own, RCT programmes cannot explain why or how these changes happen (Maxwell, 2009). This latter area is the domain of intensive descriptive studies, which in research on materials would include classroom observations of materials in use, and discussions with teachers and learners. Because of their intensive nature, such 'qualitative' studies are necessarily small in scale and therefore cannot reach general conclusions. Ideally, intensive descriptive studies should accompany RCTs in order to couple predictions for the population with an understanding of change mechanisms, and hence of their implications for policy and practice.

Policy

The research problem that animated this study focused on the relative effectiveness of workbook resources in comparison to conventional textbooks. We addressed this questions by comparing what is regarding as an example of a well designed workbook, the Primary Mathematics Research Project Intermediate Phase mathematics workbook, and comparing it to an off-the-shelf textbook, published by one of the larger commercial textbook companies, which has been in widespread use for almost two decades. We found that not only did both the Primary Mathematics and Classroom Mathematics groups make statistically significant gains between the pre-test and the post-test, but the magnitude of gains are substantially higher than most interventions in South Africa typically achieve (Taylor, 2007).

The fact that the majority of South African learners do little reading and writing has become well established over the last decade (Fleisch, 2008). What to do about this is less clear. One line of reasoning would suggest that the reason why children do not read and write is because books currently available are unsuitable. Therefore, this argument continues, we need to design and distribute workbooks, which will facilitate reading, writing and curriculum coverage. The most important conclusion of the present study is to seriously question the validity of this assumption. Our study shows that, in Gauteng at least, children cannot use books, suitable or otherwise, because they are simply not available in schools: in the 22 control schools audited, only 2 had sufficient

Grade 6 mathematics textbooks for all learners, and around half had only 1 or 0 copies. Further study is needed to examine the reasons for this situation, with possible explanations including corruption in the procurement of books, poor school management systems resulting in books not being cared for, and reluctance on the part of teachers to use books and the consequent low demand from schools. The first task in getting teachers to teach reading, writing and calculating is to get books to schools, and not necessarily to expend resources on designing new materials.

The present study shows that a representative sample of Grade 6 children from Gauteng schools in quintiles 1-4 exhibit the same degree of learning improvement in mathematics, whether they use a conventional textbook or a workbook specifically customised to address the problems exhibited by poor learners. One may argue that PMRP is itself based on a faulty assumption concerning the importance of automaticity, and that a book designed on 'outcomes-based' (or any other) principles would be more successful than either PMRP or CM. This may be a plausible hypothesis, but it is not sufficiently robust, nor supported by any evidence, to warrant spending considerable resources testing it in a full-scale national programme. It should be tested and refined in a pilot scheme before going to scale.

Although the gross learning gains exhibited by the two groups of learners are almost identical, interesting differences emerge when scores are disaggregated by learner ability and by mathematical skill. Although these details are relatively minor when compared with the overall improvement, they do have implications for materials design. As could have been predicted, the Primary Mathematics materials appear to have strong and consistently higher gains for learners in the four arithmetic operations and mental maths. This is consistent with the original intention of the designers of Primary Mathematics. As noted earlier, Schollar (2008) stressed the importance of what he termed 'bedrock' skills, which he associated with the capacity to perform mental calculations through formal learning processes or algorithms. Complementing these findings is the conclusion that *Classroom Mathematics* appears to be more successful at promoting the development of more complex concepts, including elementary fractions and problem solving strategies. The most obvious implication of these findings is that proficiency in the four operations is promoted by mental maths exercises, while problem solving skills are developed through a systematic approach throughout the course. This suggests that some combination of PM and CM would be an improvement on either book on its own.

However, we have looked at only two of the dozens of books available in South Africa, and then in a broad-brush way which glosses over the mechanisms responsible for the learning gains observed, and there is no doubt that a closer and deeper look at these books, as well as further randomised trials and qualitative studies of other materials, will enrich our understanding of the relationship between materials, pedagogy and learning. Our future investigations in this direction plan to include both the workbooks currently being designed by the DBE and the textbook used in Singapore schools, widely regarded as a key element in that country's educational miracle over the last 30 years.

7. References

Abadzi, H. (2006). *Efficient Learning for the Poor: Insights from the Frontier of Cognitive Neuroscience*. Washington: World Bank.

Borman, G., Maritza Dowling, M., & Schneck C.(2008). A multisite cluster randomized field trial of Open Court Reading. *Education Evaluation and Policy Analysis,* 30(4), *389-407.*

Bornman, G. D. (2005). National efforts to bring reform to scale in high poverty schools: Outcomes and Iimplications. *Review of Research in Education*, 29, 1-28.

Du Toit, R. (2010). Singapore Mathematics Materials Study. Joint Education Trust (manuscript).

Dugmore, C. (2008). *Grade 6 learners improve literacy skills, but maths remains huge challenge.* Press Statement by Western Cape Education MEC Cameron Dugmore.

Educational Evaluation and Policy Analysis, (2007). Introduction. 29(1).

Fleisch, B. (2008). *Primary Education in Crisis: Why South African Schoolchildren Underachieve in Reading and Mathematics*. Cape Town: Juta.

Fuller, B. & Snyder, C. (1991). Vocal teachers, silent pupils? Life in Botswana classrooms. *Comparative Education Review,* 35(2), 274-94.

Glewwe, P., Kremer, M., & Moulin S. (2007). *Many Children Left Behind? Textbooks and text scores in Kenya* Work in progress, available from www.econ.yale.edu/~egcenter/infoconf/kremer accessed April 2010.

Gronlund, N. (1998). Assembling, Administering and Evaluating the Test. In N. E. Gronlund, *Assessment of Student Achievement*. (6th Ed.) Boston: Allyn & Bacon.

Lockheed, M.E., & Verspoor, A.M. (1991). *Improving primary education in developing countries*, World Bank Publication, OUP.

Maxwell, J., (2009) Causal explanation, qualitative research, and scientific inquiry in education. *Educational Researcher,* 33(2), 3-11.

Minister of Basic Education. (2010). Budget Speech. 25 March 2010 http://www.pmg.org.za/print/20762.

Office of the Presidency. (2010). 2010 State of the Nation Address.

Raudenbush, S. (2005). Learning from attempts to improve schooling: The contribution of methodological diversity. *Educational Researcher*, 34(1), 25-31.

Scheiber, J., Brown, M., Lombard, S., Markides, M., Mbatha, K., Randall, I., &van Noort, D. (2009). *Classroom Mathematics: Grade 6*. Johannesburg: Heinemann.

Schochet, P.Z. (2008) Statistical power for random assignment evaluations of education programs. *Journal of Educational and Behavioural Statistics*, 33(1), 62-87.

Schollar, E. (2008). *Final Report of the Primary Mathematics Research Project* '2004 to 2007'. Presentation to the Conference What's Working in School Development, JET Education Services, available at www.jet.org

Schollar, E. & Kemp, A. (2003). *Back to Basics! Getting Learning Outcome One Right Intermediate Phase* (Learner Workbook and Teacher Manual).

Schreuder, B. (2008). *Improving Literacy and Mathematics in the Primary School: The Country's Most Important Priority.* Paper presented to the Conference What's Working in School Development, JET Education Services, available at www.jet.org.za

Slavin, R. & Lake, C. (2008) Effective programs in elementary mathematics: a Best-evidence synthesis. *Review of Educational Research* 78(3), 427-516.

Slavin, R. (2008). What Works? Issues in synthesizing educational program evaluations. *Educational Researcher,* 37(1), 5-14**.**

Taylor, N. (2007). Equity, Efficiency and Development in South African Schools. In T Townsend (ed.) *International Handbook of School Effectiveness and Improvement*. Dorchrecht, Netherlands: Springer.

Van der Berg, S., & Louw, M.(2008). South African student performance in regional context. (2008). In Chisholm, L., Fleisch, B., and Bloch, G., *Investment Choices for South African Education.* Johannesburg: Wits University Press.

NOTES

[1] This project is a joint undertaking of colleagues at the Wits School of Education and the JET Education Services. Funding for the study has been provided by the DG Murray Trust and JET. We would like to acknowledge a number of individuals who have contributed their time and expertise. Linda Zuze has played an invaluable role in helping determine the sampling frame and providing statistical training. Helen Perry and Prof. Peter Fridjhon assisted with some of the data analysis. Talia de Chaisemartin has provided guidance in the analysis of the qualitative data. Special thanks to Andres Martinez at the University of Michigan's Building Capacity to Evaluate Group-Level Interventions and Stephen Raudenbusch at the University of Chicago for assistance with study design and sample size. We would also like to thank the Gauteng Department of Education for permission to undertake the study and the Non-Medical Ethics Committee of the University of the Witwatersrand reviewed the proposal and issued an ethics clearance certificate for the study. The findings, interpretation and conclusion expressed in this paper however are entirely those of the authors. We welcome comments and criticism to Brahm.Fleisch@wits.ac.za

[2] The full programme starts at grade 4 level and ends at grade 6 level. Therefore, an evaluation should track the Grade 4 class to Grade 6 to allow the effect of the intervention to be measured.

[3] The power of the statistical test refers to the probability of avoiding a Type II error (i.e. incorrectly rejecting a null hypothesis). Therefore it represents the likelihood of drawing the correct conclusions about the significance of differences between groups. Typically, a power level of 80% is considered high enough to detect differences while keeping sample sizes reasonable.

[4] The ICC is the level of school variation as a proportion of total variation (student and schools). It can also be described as the level of inequality between schools. The higher the ICC, the larger are the systematic differences in achievement scores between schools and the more groups required in the sample. Because of the skewed patterns of achievement across the South African schooling system, we will exclude the wealthiest 20% of schools from the sample to ensure an ICC that falls within these bounds.

[5] In order to determine appropriate sample size, it is necessary to have some prior knowledge of expected size of the intervention effect. In much of the contemporary US based literature this is has been standardized to a common effect size unit, i.e. percentage of the standard deviation of the outcome measure. This allows for comparison across studies using different scales. While the original PRMP study did not report results in percentage of the standard deviation of the outcome measures, the percentage point gains reported were very high. The use of .33 of a standard deviation is a conservative measure within the bounds of high effects in US based interventions.

[6] Some school sizes decreased substantially since selection to actual data collection.

[7] In the original pilot, we had 22 schools in each group. In the final study in 2010, one school from each group dropped in the study (mainly due to pressure from district office to use the Foundations for Learning plans rather than the lesson plans in the study.)