THE EDUCATION TRIPLE COCKTAIL

System-wide instructional reform in South Africa

BRAHM FLEISCH

THE EDUCATION TRIPLE COCKTAIL

System-wide instructional reform in South Africa

Brahm Fleisch

The Education Triple Cocktail:
System-wide instructional reform in South Africa

First edition 2018

Published by UCT Press, an imprint of
Juta and Company (Pty) Ltd
PO Box 14373, Lansdowne 7779, Cape Town, South Africa
www.uctpress.co.za

ISBN 978 1 77582 246 2 (print)
ISBN 978 1 77582 247 9 (Web PDF)

Project manager: Mmakasa Ramoshaba
Editor: Language Mechanics
Proofreader: Renee Moodie
Cover designer: Marius Roux
Typesetter: Nazli Jacobs
Indexer: Tessa Botha

Typeset in 11pt on 14pt Minion Pro

This book has been independently peer-reviewed by academics who are experts in the field.

Table of Contents

Acknowledgements

The ideas and research contained in this volume have been a long time in the making. Much of the first draft of the manuscript was written during my 2011 sabbatical leave. I need to acknowledge the University of the Witwatersrand for providing me with the time and space to write. Although they were not always in agreement with my thinking, my colleagues and students have always pushed me to clarify my ideas. In particular, I would like to thank Stephen Taylor, Nompumelelo Mohohlwane, Janeli Kotze, Francine de Clercq, Yael Shalem, Linda Chisholm, Debbie Botha, Zakhele Mbokazi, Nic Spaull, Kerryn Dixon, Volker Schöer, Mary Metcalfe, Gareth Roberts, Godwin Khoza, Dhianaraj Chetty, Hamsa Vinkat, Ingrid Sapire, Lynne Slonimsky, Hannchen Koornhof, Tom Pugh and Kaitlin Alsofrom. I would also like to thank friends and colleagues Andy Hargreaves and Dennis Shirley for the many conversations over the past ten years on the Wild Coast hikes, at Walden Pond and Boston College. My greatest debt is to my many colleagues in schools with whom I have worked on the Gauteng Primary Language and Mathematics Strategy, the Reading Catch-up Programme and the Early Grade Reading Studies. I need to make special mention of and thank two extraordinary educators, Lorraine Marneweck and Penny Groome, and their team for the huge contribution to the advancement of our understanding of how to improve instruction at scale. The National Research Foundation contributed to the funding for the book. Finally, I would like to thank the team at University of Cape Town Press and Juta for helping me with the preparation of the manuscript. Acknowledgement is due to PIRLS for permission to use material that has appeared in a different form elsewhere.

Profile

BRAHM FLEISCH is Professor of Education Policy in the Division of Education Leadership, Policy and Skills at the University of the Witwatersrand, Johannesburg. Brahm did his graduate studies at Teachers College, Columbia University in New York. After moving back to South Africa in 1990, he has lectured in education at the Wits School of Education and served as a district director in the newly formed Gauteng Department of Education. He developed and co-ordinated the Professional Certificate in Education Finance, Planning and Education Economics. His books include: *Primary Education in Crisis: Why South African Schoolchildren Underachieve in Reading and Mathematics* and (co-authored with Stu Woolman) *The Constitution in the Classroom: Law and Education in South Africa, 1994–2008*. His current research and professional work focus on successful system-wide instructional reform. Since 2010, Brahm has served as the advisor of the Gauteng Department of Education and the National Education Collaboration Trust. He served as counsellor on the South African Council of Educators from 2012–2017. He is currently co-principal investigator on two large randomised control trial studies on instructional change.

Chapter 1
Introduction

Since AIDS was first identified in 1981, more than six million people had died from the disease by the start of the 1996 Vancouver AIDS conference. On the last day of the meeting, Dr Martin Markowitz, a researcher from the Aaron Diamond AIDS Research Center, came to the podium to describe the early results of a small clinical trial. Twelve young adult male patients who had recently been diagnosed were given a combination drug therapy that included AZT and 3TC with a new type of drug called a protease inhibitor. AZT and 3TC treated the early stages of the disease; the protease inhibitor was designed to block the multiplication of an enzyme critical to the spread of the virus in the body. Dr Markowitz described how within nine months of receiving the new 'triple cocktail' treatment, HIV levels in 9 of the 12 men had fallen below detectable levels and their white cell counts had risen sharply.

The triple cocktail is not a cure, but in time it changed AIDS from the leading cause of death of young adults to a manageable condition. The discovery of the efficacy of the triple-drug regime was a turning point in the treatment of the global epidemic. That said, it took more than a decade for the proven treatment to reach the majority of patients in need. In the early stages of the use of the triple cocktail, the cost of treating a single patient was up to US$20 000 per year, unaffordable to all but the richest people and wealthiest countries. What complicated the matter was that it was not a simple treatment; to deliver it at scale required an advanced health system.

Over the next 15 years, the collective effort of hundreds of researchers, health professionals, activists and government officials in the Global South would make treatment available to millions of patients in resource-limited countries. The Durban AIDS conference in 2000 pivoted on debates about treatment access for the world's poorest countries. An implementation breakthrough came in 2003 when researchers discovered how to reduce the number of pills required to be taken daily from around 15 to 2, and activists and governments negotiated the use of generic drugs that brought down the price to US$150 per person per year. Even with the cheaper and simpler-to-administer triple cocktail, many HIV patients still die, their deaths caused by severe malnutrition, late-stage take-up of the treatment, tuberculosis when treatment starts and anaemia (Vella et al, 2012).

Although seldom in the newspaper headlines these days, HIV and AIDS, the disease and the universal rollout of treatment, have touched most South Africans. It is an extraordinary story of success against what appeared to be overwhelming adversity, with the scientific discovery of an efficacious treatment and struggles for cost-effective ways to scale up treatment at its centre. It was the collective efforts of researchers, funders, bureaucrats, activists, lawyers and front-line health professionals that turned the disease from a post-apartheid national catastrophe into a difficult but manageable healthcare disease.

This book draws inspiration from this AIDS treatment story. The obvious way is indicated by the name given to the early grade learning intervention model, the Education Triple Cocktail, which is at the centre of the latter chapters of the book. But more than that, the global partnerships and international collaborations, the central role of empirical research, the urgency and levels of the funding provided, the recognition of the unique challenges of rolling out access in resource-limited settings and the solutions that were found to all of these challenges offer valuable lessons and inspiration to education researchers, activists and government officials striving to ensure that all children in school become successful learners, not only in South Africa, but across the Global South.

In the concluding chapter of my previous book, *Primary Education in Crisis*, I argued that what was needed to address the epidemic of underachievement in primary school reading and mathematics was rigorous evidence of 'what works'. I had been convinced by the writings of a University of Chicago economist, Stephen Raudenbush, who argued that evidence-based instructional programmes, what he referred to as 'instructional regimes', could be used (like antiretroviral therapy regimens with AIDS) to fundamentally alter the patterns of instruction and resultant performance of primary education systems. Notwithstanding the pitfalls of 'policy borrowing' and adopting health sciences discourses for the study of educational problems, I was (and continue to be) convinced that a great deal of value can be extracted from evidence gained using scientific methods.

With this as the backdrop, my research agenda over the past decade has focused on helping to build such a rigorous knowledge base. While trained as an historian, I have moved over to what some of my colleagues used to call the 'the dark side' — quantitative research, and specifically the use of experimental and quasi-experimental methods to advance our understanding of how to improve learning outcomes at scale for poor and working-class learners, and narrow the unequal outcomes of our school system.

The first steps in the journey came with my involvement in a randomised control trial study designed to re-test the efficacy of a promising Intermediate Phase school mathematics instructional programme that used structured and differentiated worksheets in a set of self-contained workbooks. With the help and encouragement of the late Peter Fridjon, Dr Nick Taylor, Ingrid Sapire and Roulien

Herholdt, I (Fleisch et al, 2011) designed our large-scale field experiment. The results were not what we anticipated — while the learners in the treatment 'workbook' group of schools did better than those in the 'control' group who got a standard textbook, and both groups improved substantially on the endline test, the intervention unit costs were substantially higher than those for the standard textbook. Our first field experiment demonstrated not only that cluster randomised control trials research was possible, but because of the rigorous design features, the findings would be powerful.

Advances in knowledge sometimes come from fortuitous occurrences. In 2010 I was asked to prepare a provincial strategic plan for an early grade reading programme to be implemented by the Gauteng Department of Education (GDE). Over the next four years, I advised the GDE management team, including Len Davids, Deb Botha, John Thurlow, Kholo Makhaga, Ingrid Sapire, Elspeth Khembo and Jonathan Williams as they developed a system-wide intervention model to improve early grade reading and later mathematics. It was during my research on the Gauteng Primary Language and Mathematics Strategy (GPLMS) that my thinking about change at the instructional core (Richard Elmore's phrase) and system-wide reform models crystallised. I conducted qualitative case studies of various aspects of the intervention model, examining how teachers talked about the prescriptive lesson plans, their experiences working with instructional coaches in their classrooms, and asking questions about the new educational materials that were part of the programme. Unfortunately, a planned external evaluation of the GPLMS never happened, and though many colleagues in the sector who came and saw the instructional work in the schools were impressed, there was little rigorous evidence of the efficacy of the education triple cocktail model.

A group comprising Volker Schöer, Gareth Roberts and Amy Thornton from the Wits University School of Economics and I got started on the question of impact through the use of a hybrid regression discontinuity design/natural experiment (Fleisch, Schöer, Roberts & Thornton, 2016). The findings from this study were positive and promising, albeit limited to a local area treatment effect (LATE). Excited by the parallel experimental research coming out of Kenya and India, we suspected that our research on system-wide instructional reform was moving in the right direction.

What was needed was more and better evidence about whether the education triple cocktail model worked. Specifically, we needed experimental evidence of its efficacy at varying dosages, when implemented across a range of schools, when delivered in a rural setting, with different school subjects and at different grades or phases. To answer the dosage and phase questions, an expanded team that consisted of Dr Stephen Taylor from the Department of Basic Education, Dr Thabo Mabogoane from the Department of Planning, Monitoring and Evaluation, Volker Schöer and I designed an experiment to test a system-wide remedial version of the GPLMS

triple cocktail model. This study was conducted in 2014 in the Pinetown district of the KwaZulu-Natal Department of Education. The results of this study (Fleisch, Taylor, Schöer & Mabogoane, 2017a) showed that the intervention's overall impact on learner performance was statistically equivalent to the gains made by learners in the control schools. Although alternative explanations were explored, the likely reasons for the limited impact were that the intervention offered 'too little, too late', the length of the intervention was too short and that it was difficult to catch up on missed learning by the time children had reached the Intermediate Phase.

In 2015, the Early Grade Reading Study (EGRS), the most ambitious experimental research programme ever undertaken in South Africa, was initiated and led by Dr Stephen Taylor, director of research in the Department of Basic Education. As part of this ambitious venture, the study designers used the experiment to test the impact of the triple cocktail model, comparing its impact to a comparable control group, a second group that received a similar model that involved training rather than coaching as well as a parent training model. The EGRS I and thereafter the EGRS II (English as a First Additional Language) provided the compelling evidence we needed to plan system-wide improvement. The team at the Department of Basic Education, Stephen Taylor, Nompumelelo Mohohlwane and Janeli Kotze, need to be acknowledged as pioneers, showing the way in designing and executing rigorous research to solve pressing education problems. The concluding chapters of this book summarise the exciting cutting-edge findings.

Chapters 2 to 4 set up the 'learning' problem and outline the conceptual framework for change at the instructional core. I begin by rehearsing the key research undertaken in the past decade on the patterns and prevalence of primary school underachievement, and identify new features of South Africa's bimodal learner performance. Specifically, Chapter 2 points to the limited progress that has been made, the substantial difference between reading in the Home Language and the First Additional Language, and insights from a new generation of research studies indicating an underlying decoding weakness that anticipates the poor reading comprehension. While there are questions that can be asked about both the cross-national achievement tests and the national evaluation studies, when all the studies are interpreted as a whole, a troubling picture emerges.

Chapter 3 reviews government's attempt to use various national policies to improve learning outcomes. These include policy interventions aimed at teacher capacity-building, curriculum reform, institutional accountability, financial incentives and the provision of educational materials inputs. While the South African education system has stabilised over the past two decades, the chapter explores why the many well-intentioned policies have had only limited impact on entrenched instructional practices and, by extension, failed to narrow the achievement gap between fee and no-fee schools. But why has policy failed? The chapter develops an argument that the problem with national policy as an agent of change is that its

policy planners fail to find the right lever to nudge (or rupture) the system's 'instructional core'. This idea is at the centre of this book — the recognition that the core of the education system is not made up of distinct components such as teachers, textbooks and time, but a combination of these to make up 'instructional practices' that manifest into millions of lessons, activities and tasks involving teachers and learners around content while using materials. The conceptual pivot towards an analytical 'instructional practice' as the proximal cause of learning reveals the limitation of the two main change theories, that is, input/outcome and school/teacher effectiveness.

What exactly is instructional practice? How does it manifest in the South African context? Drawing on little-known but highly innovative qualitative research studies, Chapter 4 paints a composite portrait of two distinct types of 'instructional practices' to be found in early grade classrooms in South African schools. The first is the instructional practice that is predominant in most 'no-fee' schools in townships and rural schools around the country. Although often associated with 'rote learning' or the 'rote rhythm method' (Macdonald, 2002), it is actually a much more complex practice. The practice is evident in the early grade teaching of reading in African languages. It often begins with the teaching of letter names and a syllabic approach (ie *ma-me-mi-mo-mu*) and moves on to letter sounds, simple words and sentences. Reading books are introduced only in the second and third year of schooling. Few texts are taken home and there are almost no opportunities for young children to experience independent reading of children's storybooks. While 'big books' have been added to this instructional practice, they are more often than not used for rote whole-class chorus reading. Most of the actual reading activities involve choral or chorus reading aloud, copying from the board and filling in worksheets.

Many of the studies reviewed point to the pervasiveness of the practice of choral reading aloud, not only in the Foundation Phase but also in the Intermediate Phase. If, as Elmore asserts, instructional practice predicts academic performance, then children would do well in a test that asked them to repeat a word or sentence that has been spoken orally. Children with little experience and few skills in independent reading and meaning-making with texts would struggle with test questions requiring these competencies. There is an important exception in working-class and rural schools, pertaining to the reading skills of a small group commonly referred to as the 'clevers', or precocious readers. These children are singled out to become the public readers, children who would be 'readers' in public events or for outside visitors.

The other type of instructional practice is predominant in former white, Indian and many coloured schools. The early grade instructional practices associated with teaching literacy in these schools involve the combination of core methodologies, such as synthetic phonics, teaching high-frequency words using the 'look and say

method', building fluency using decodable texts, basal or levelled reading series, and opportunities to do independent reading. The graded or levelled reading schemes, such as the Ginn New Reader 360 or Oxford Reading Tree, are often the core educational materials central to this instructional practice. These schemes or series consist of sets of small books organised from simple to complex, and designed to match the reading levels of individuals or groups within a classroom. Children move up levels (colours or numerical value from 2.1 to 2.2) over the duration of a year of instruction or even across grades. Along with these series, teachers in these schools use commercial phonics programmes such as *Letterland* and *Jolly Phonics*. Foundation Phase teachers in these schools have been trained as early grade teachers, and are familiar both with the theory and use of the techniques of teaching early grade literacy.

Over the past half century, a variety of innovations have become mainstreamed in this instructional practice, including the use of shared reading techniques and authentic texts for independent reading. Home-based support plays a key role in helping to support young children to develop fluency. This instructional practice requires class teachers to assess all children on a regular basis, assign them to reading groups, identify the most appropriate levelled readers and monitor progress.

During the last four decades there have been many initiatives, both government and NGO, that have aimed to transform instructional practice and improve reading outcomes in township and rural primary schools. Chapter 5 analyses the lessons learned from these efforts, drawn from published accounts. One of the key insights from the review is how little each successive intervention improved on the successes and the failures of the ones that came before. While there is a history of innovation in reading programmes and initiatives, there was little real effort to rigorously evaluate what worked, or to build on and advance promising practices.

At this point in the book's narrative, there is a shift in gear. The second half of the book provides the background, the theory and the empirical evidence of a new approach to change at the instructional core. This new approach, called a structured pedagogic programme in the international literature (Snilstveit et al, 2016), has come to be called the education triple cocktail in South Africa. This theory is confirmed by empirical evidence from experimental research and large-scale implementation that represent knowledge advances in the field of educational change. The set of propositions or hypotheses that set out to explain the mechanics of change are tested by actual data from the real world.

Drawing on change theory generated from leading thinkers in the field and change experiences around the world, Chapter 6 outlines the theory of action that guided the original intervention, which was designed to change the instructional core of early grade reading in township primary schools. It was this theory of action that animated the series of field experiments in the Reading Catch-Up Study and the Early Grade Reading Study over the last five years. Chapter 7 goes into detail about the first effort to rigorously evaluate the triple cocktail theory of change. As

with any experimental programme, setbacks become the springboards.

Chapter 8 summarises the compelling evidence for the educational triple cocktail, beginning with the results from the regression discontinuity design/ natural experiment in Gauteng and the comparative interprovincial verification, the Annual National Assessment statistics in 2014, and ending with the results of the Early Grade Reading Study in North West and Mpumalanga.

The contribution to the field of educational change can be summarised in the following propositions:

- There needs to be a strong alignment between official curriculum and the new instructional core.
- External standardised assessment systems are not essential, especially in an environment of low trust.
- The focus of change must be on instructional practice itself, not on a signifier or indicator of instruction.
- When working with the new instructional core, it must be understood as a set of learning activities and tasks that require teachers to take on new routines, new skills and new knowledge, and that it requires the provision of new and appropriate educational materials.
- The old and new instructional core is observable in the activities and tasks children actually do.
- The activities and tasks together constitute the instructional practice. It is this practice (collection of activities and tasks) that drives improved academic performance.
- Change at the instructional core is first about putting in place new routines or daily practices. In most cases, shifts in teacher beliefs and knowledge follow.
- School leadership (the role of heads of departments and principals), district support and monitoring and ongoing improvement are critical for institutionalising and sustaining the new instructional core.
- The educational materials are driven by a coherent and evidence-informed new instructional core. In the case of Foundation Phase literacy, this includes coherently and consistently supplying educational materials for all aspects of literacy development, ie graded readers, phonics programmes, books for independent reading and workbooks for writing.
- In order to assist teachers in the change processes, the educational materials need to be prescribed.
- The learner materials have two distinct purposes: they provide the basis for activities and the tasks, but they also are key to the teachers who are learning the new routines and new core methodologies.
- Teachers beginning the change journey with the new educational materials need just-in-time, lesson-specific training. Once on the journey, teachers also need to

understand the big picture, the theory that sits behind the new instructional core.

- Along with just-in-time training, teachers need one-on-one instructional coaching. The purpose of coaching is to provide guidance and direction on the new practice, but coaching has an equally important role in providing emotional support as teachers leave their old practice behind and begin to adapt to the routines and rhythms of the new practice.
- As interventions move towards system-wide implementation, there is a need to standardise the practice of the coaching itself.
- As the new instructional practice begins to show signs of success, teaching becomes more rather than less demanding. This is because the gap between stronger learners and weaker learners widens, and teachers have to contend with this widening in class differences.
- Once the technical aspects of the new instructional core begin to settle, the secondary systems that reinforce the older system of instruction need to be activated.
- These include how schools work with parents; how educational materials are managed and budgeted for; how teachers assess individual learners; the tools for such assessment and the record-keeping; and, more broadly within the school, the human-resource planning and class-size management.
- System managers need to recognise and plan for differentiated teacher responses to change at the instructional core, and plan strategies for early adopters, the compliant middle and the active resisters. The instructional core is dependent on teachers in the compliant middle institutionalising the new practice.

Change at the instructional core must begin by conceptualising a new core instructional practice. In this context, the new instructional practice largely draws on a dominant national and international instructional practice, which has both a strong base in terms of academic knowledge and extensive examples of effective practice widely distributed around the country. While it may be that the construction of the new instructional core is not a single event, it is central to the change process. The new instructional core then needs to be built with resources, materials or learning tools.

Chapter 2
Not achieved

Introduction

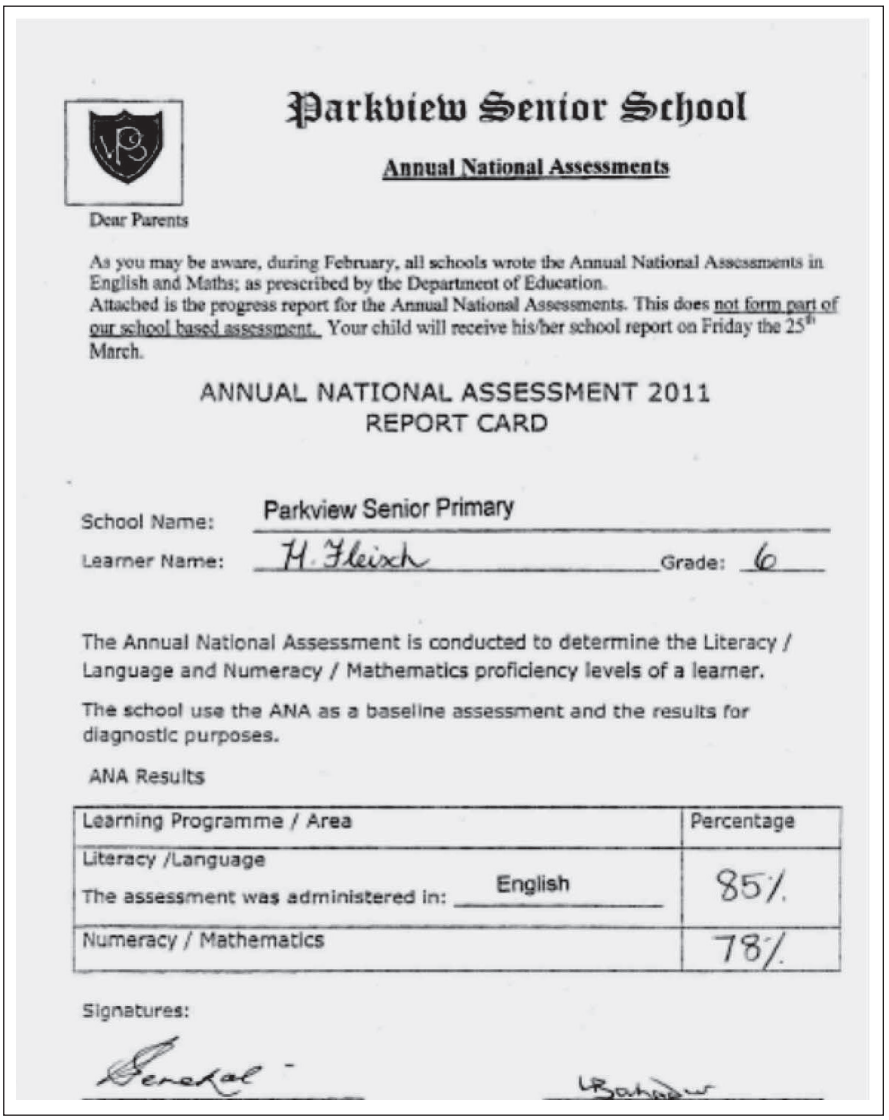
Shocking as they are, no one should be surprised by the results. That is how the Minister of Basic Education began her press conference on the 2011 Annual National Assessments (ANA) in the beautiful briefing room at the Union Buildings, Pretoria. While administrative difficulties had delayed the results of the universal testing of all children in public schools in Grades 1 to 6, the verification ANA, a subsample of the test that was moderated by an external agency, showed that the vast majority of South African schoolchildren in February 2011 were not even close to achieving the minimum proficiency in reading and mathematics. On these two core subjects that form the foundations of later learning, reading and writing in the language of learning and teaching (LOLT) and numeracy/mathematics, the overwhelming majority of South African primary-school children scored below 35 per cent. Therefore, what the Department of Basic Education described as the benchmark level was *not achieved*. Not achieved!

Simply put, this book is about why the majority of South African primary-school children are unable to read at the levels required by the official curriculum and how we can use evidence to fix this problem.

The ANA reports (like the one in Figure 2.1) that millions of children took home in 2011 communicated to parents that their children were not at the levels that the Department of Basic Education expected. While some parents might have blamed their children, in truth it was the education system — the teachers, the district staff, the policy-makers and the politicians — who were and are failing our children. Unlike my previous book, which focuses on poverty and other contextual factors, this book begins and ends in the classroom. My objective is to explore the intimate relationship between teaching and learning failure.

New evidence of the bimodal pattern of achievement

I begin with the evidence for this school system failure. At the time I completed *Primary Education in Crisis*, there was already an abundance of research evidence that consistently pointed to two striking insights. First, South African primary-school children are systematically underachieving in reading and mathematics.



Parkview Senior School

Annual National Assessments

Dear Parents

As you may be aware, during February, all schools wrote the Annual National Assessments in English and Maths; as prescribed by the Department of Education.
Attached is the progress report for the Annual National Assessments. This does not form part of our school based assessment. Your child will receive his/her school report on Friday the 25th March.

ANNUAL NATIONAL ASSESSMENT 2011
REPORT CARD

School Name: Parkview Senior Primary

Learner Name: H. Fleisch Grade: 6

The Annual National Assessment is conducted to determine the Literacy / Language and Numeracy / Mathematics proficiency levels of a learner.

The school use the ANA as a baseline assessment and the results for diagnostic purposes.

ANA Results

Learning Programme / Area	Percentage
Literacy /Language The assessment was administered in: English	85%
Numeracy / Mathematics	78%

Signatures:

Figure 2.1: Example of School Report Card for the Annual National Assessments

Second, the pattern of failure closely tracks the wider pattern of inequality in our society. Cross-national studies of educational performance conducted in South Africa from the mid-1990s to the mid-2000s, government's systemic evaluation reports, large-scale base-line surveys, and dozens of small-scale qualitative studies, consistently confirmed this reality. The evidence showed that between 60 per cent and 70 per cent of South African schoolchildren were getting to the end of primary school without being able to read, write and do mathematics at a proficiency level

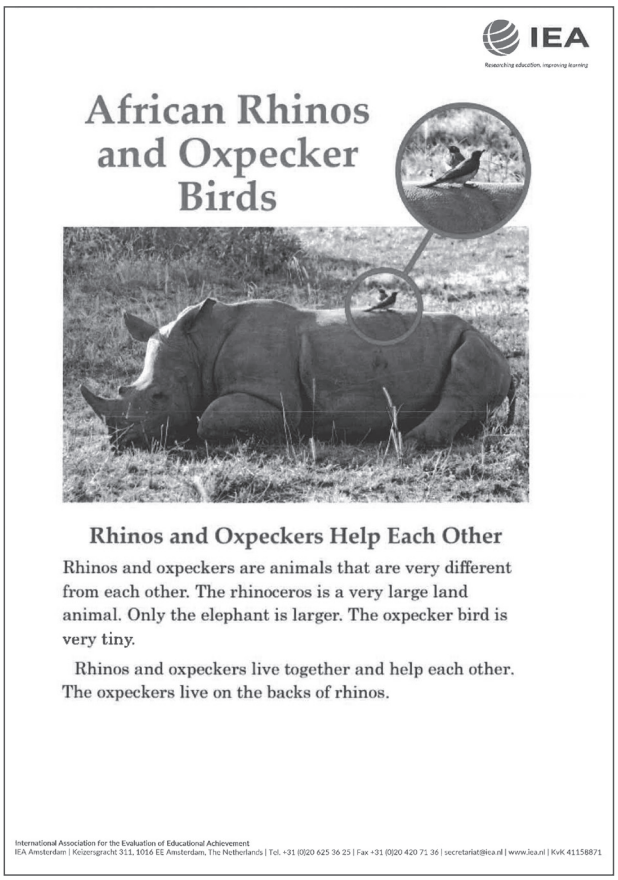
required by the official curriculum. In concrete terms, the research showed that our primary school leavers were unable to read a tabloid newspaper article fluently, or locate and communicate basic explicit information contained in it. Most primary-school children struggled with basic arithmetic operations, many using inappropriate and ineffective methods to solve simple problems. And while cross-country comparisons placed South African at the bottom of the system league tables, the national scores masked the underlying pattern of inequality. The majority of schoolchildren from predominantly black and coloured working-class and poor communities were substantially behind their counterparts attending schools in more affluent communities.

What is new in the research about primary-school children academic achievement? First, the evidence shows that little has changed over the past decade. The best indicator comes from the Progress in International Reading Literacy Study (PIRLS) 2011 and 2016, which shows that the same small proportion of Grade 4 learners is reaching the low international reading benchmark. Second, in addition to the fact that a substantial proportion of schoolchildren cannot read fluently and understand what they read in their Home Language, we now know that the problem is even more acute for reading in English when we test children who began learning in an African language and shifted to reading in English in Grade 4. Third, we are developing a finer-grained picture of the prevalence and patterns of illiteracy. Finally, until recently, researchers tended to report reading achievement as an aggregate score composed of various subskills particularly focused on comprehension. Over the past 10 years, the research community has become more attentive to the specific components of reading and is moving beyond the single composite indicator. The Early Grade Reading Assessment (EGRA) has found its way into a number of different testing processes, from the National Education Evaluation and Development Unit (NEEDU) 2013 report to testing in randomised control trials, such as the Early Grade Reading Study. The increasing use of EGRA or EGRA-like instruments is beginning to help unpack differentiated mastery of reading subskills. Most clearly evident is the emergence of attention to reading fluency, as measured by Words Correct Per Minute (WCPM) as a valid and reliable proxy for reading.

Reading progress and the PIRLS 2006, 2011 and 2016

The Progress in International Reading Literacy Study (PIRLS) is one of a handful of international studies of academic achievement. Unlike its more famous cousins, Themes in International Mathematics and Science Study (TIMSS) and Programme for International Student Assessment (PISA), the PIRLS study measures reading literacy, reading behaviour and attitudes of Grade 4 and 5 learners.

What is particularly important for South Africa is that the PIRLS test results provide a picture of how South African learners have or have not progressed over time. In 2006, 16 073 Grade 4 learners were tested in all 11 languages, and 15 657 Grade 5 learners in English and Afrikaans. Five years later, in 2011, a similar-size sample of Grade 4s were tested in all languages, and a smaller sample of Grade 5 learners were tested in English and African languages using what was then referred to as a pre-PIRLS test. In 2016, representative samples of Grade 4 and Grade 5 learners were assessed again using passages from both the PIRLS literacy (formerly referred to as pre-PIRLS) instrument and main PIRLS test. Learners were tested in all of the country's 11 languages in Grade 4 and in English, Afrikaans and isiZulu in Grade 5.



IEA

Researching education, improving learning

African Rhinos and Oxpecker Birds

Rhinos and Oxpeckers Help Each Other

Rhinos and oxpeckers are animals that are very different from each other. The rhinoceros is a very large land animal. Only the elephant is larger. The oxpecker bird is very tiny.

Rhinos and oxpeckers live together and help each other. The oxpeckers live on the backs of rhinos.

International Association for the Evaluation of Educational Achievement
IEA Amsterdam | Keizersgracht 311, 1016 EE Amsterdam, The Netherlands | Tel. +31 (0)20 625 36 25 | Fax +31 (0)20 420 71 36 | secretariat@iea.nl | www.iea.nl | KvK 41158871

Figure 2.2: Example of an informational passage

Source: http://timssandpirls.bc.edu/pirls2016/international-results/wp-content/uploads/structure/PIRLS/11.-appendices/H_restricted-use-passages-questions-and-scoring-guides.pdf. (Thanks to Nic Spaull for pointing to the Rhino/Oxpecker passage.)

Consistent across all three time points, the objective of the PIRLS process was to assess how well South African schoolchildren can read at multiple levels, and how they compare to their international counterparts. Although two different instruments have been used, that is, PIRLS and pre-PIRLS (literacy), with comparable passages in all tests, it is possible to compare scores between years and across systems.

The PIRLS test itself consists of four passages, two fiction and two non-fiction (informational), each followed by a set of 13–15 questions. In the South African context, the language that children were tested in Grade 4 was the language of teaching and learning used in the Foundation Phase at that school.

The 'Rhinos and Oxpeckers Help Each Other' passage is an example of an informational passage. If a learner could correctly answer questions such as 'Where do you find oxpeckers?' they could reach the low-benchmark reading level.

By analysing the patterns of correct answers to questions based on these types of passages the PIRLS researchers could determine the proportion of school-children in each country that could read at various levels. They could also compare country benchmark patterns. The PIRLS test had five levels: (1) children who could not reach the minimum reading level; (2) children who could reach the minimum reading level or low international benchmark; and levels (3), (4) and (5), which were intermediate, high and advanced reading levels. In the PIRLS system, to reach the low reading benchmark, learners would at least be able to retrieve basic information from the text to answer simple questions as in the passage. This is often referred to as being able to read for meaning. At the higher levels (intermediate, high and advanced), learners would be expected to demonstrate that they could interpret texts, see logical links, see connections between events, generalise, read and interpret tables, and at the top level, understand the writer's stance or perspective.

As with many of the international test instruments, the PIRLS test was scaled on a range from 0 to 1 000 with a central point at 500 (international mean). In the 2016 South Africa test, 23 per cent of the learners tested wrote the test in English, 22 per cent in isiZulu, 16 per cent in isiXhosa, just over 9 per cent in Sepedi and Afrikaans, with smaller percentages in the other languages.

On average, for all schoolchildren tested, the South African Grade 4 learners scored 320, substantially below the international average of 500 and below averages for the two other African countries that participated, that is, Egypt (330) and Morocco (358). Grade 4 learners who wrote the test in English (irrespective of their actual home language) scored 372, with children who wrote in African languages such as isiXhosa (283) and Sepedi (273) scoring lower. The PIRLS 2016 test also confirmed a large gender gap, with girls (347) scoring 52 points higher than boys (295).

What shocked the nation was that the vast majority of this representative sample of South African Grade 4 reading scores fell below the low benchmark level.

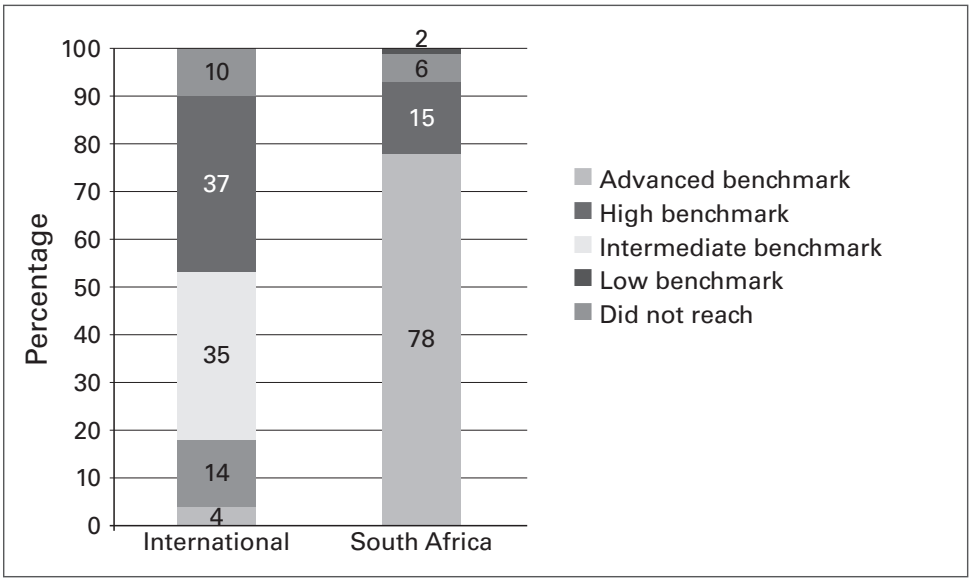


Figure 2.3: Comparison of international and South Africa benchmark attainment on the PIRLS 2016 Grade 4 test

Source: Adapted from Howie et al (2017)

Seventy-eight per cent of South African Grade 4 children were not able to reach the lowest benchmark compared to 4 per cent internationally (Howie et al, 2017:4). The assumption was that schoolchildren who did not reach the low benchmark could not really find explicit information in the fiction or informational passage.

The South African average performance on each of the PIRLS benchmark levels obscures the major disparity in achievement within the education system. With regard to learners who wrote the test in English, 43 per cent reached or exceeded the lowest benchmark, compared to only 7 per cent of learners who wrote the test in Sepedi. These statistics clearly reflect the enduring bimodal pattern of achievement within the education system.

After three rounds of testing, PIRLS can provide important insights about trends over time in South Africa. Although the results from the African language test in 2006 could not be used because the mean scores were too low, the PIRLS 2011 and 2016 tests could be compared for all languages, and a comparison is possible in English and Afrikaans scores for all three time points. An analysis revealed that there has been no real change in reading performance of Grade 4s between 2011 and 2016 tests in all languages, and essentially no change in reading performance in English and Afrikaans reading from 2006 to 2016 (there is a change if English and African scores are combined.). The proportion of learners who do not reach the low benchmark has remained more or less constant between

2011 (76 per cent) and 2016 (78 per cent), although at the top end there has definitely been a decline in the proportion of children at the high and advanced benchmark levels.

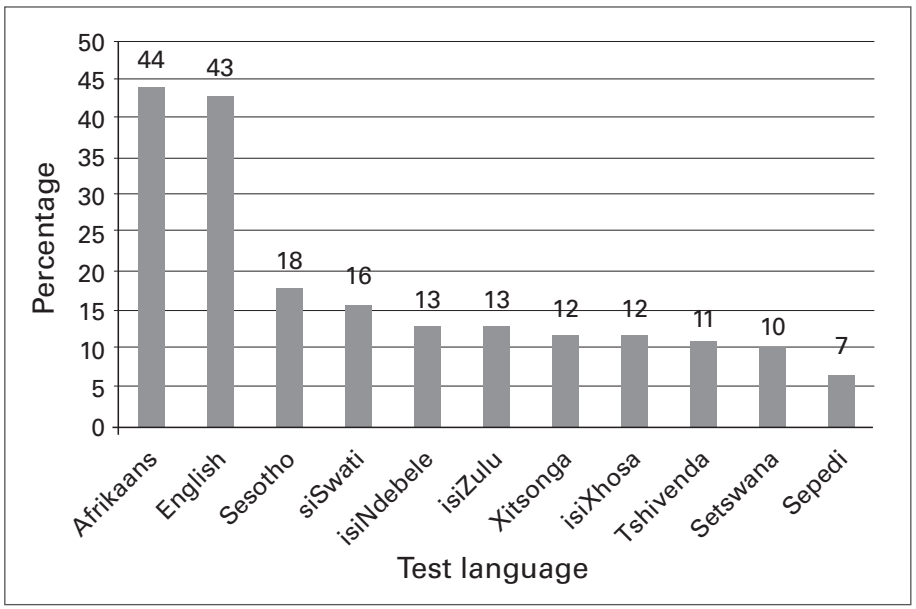


Figure 2.4: Percentage of learners who reached the Low Benchmark or above by test language
Source: Adapted from Howie et al (2017)

Reports of the National Education Evaluation and Development Unit

While international tests provide useful insights about relative performance using the broadest brush strokes, what picture emerges from within South Africa's own assessment of schoolchildren's learning? In this section, we look at the efforts of government itself to assess learner performance, both in the National Education Evaluation and Development Unit (NEEDU) reports and the Annual National Assessments that were undertaken from 2011 to 2014.

As part of the 2012 NEEDU report on the state of literacy teaching and learning in the Foundation Phase, the government team collected information on how well children were reading in the schools they visited. They did this by asking teachers to identify the three best readers in their classes and tested each child's reading ability. The team made use of an adapted version of the Early Grade Reading Assessment (EGRA), an instrument that was gaining in popularity in the assessment of reading in the Foundation Phase in South Africa and across the continent. The test has a number of subtest components designed to assess children's

competence with letter names and sounds, word reading, nonsense word reading (phonics), sentence and paragraph reading and comprehension. While the EGRA instruments had been around for almost a decade, the NEEDU report acknowledged that a number of the subtests, particularly oral reading fluency, had not been standardised in the African languages in the South African context. And there was a recognition of the unique linguistic characters of South African languages. Nguni languages such as isiZulu are agglomerative languages characterised by long words.

Notwithstanding these qualifiers, the testing of how many words children can read correctly and testing their comprehension provides a useful indication of how well children are mastering the skill of reading. Six hundred and forty-one learners (the best three in each class) were tested. And the results: 'The most striking feature … is that 72 per cent of the three best learners in each class observed are reading below the average benchmark for Grade 2 learners, and that 22 per cent were below the poor benchmark' (NEEDU, 2012:40). This was measured against an international benchmark cited by the NEEDU report that suggests that top learners should be reading on average at 125 words per minute and average learners at about 70 words per minute (wpm). The report concludes that 'if learners are not reading independently at around 50 wpm by the end of Grade 1 they are likely to struggle for the remainder of their time at school' (NEEDU, 2012:40).

The following year, the 2013 NEEDU report focused on teaching and learning, specifically in rural primary schools. Building on the 2012 research on oral reading fluency, the NEEDU researchers observed: 'There is no culture of listening to learners reading, surely the most obvious way of tracking progress of individual learners'. In the 2013 study, the NEEDU team was far more sophisticated than the previous year both in the sampling of learners and in the test instruments. Unlike the earlier study, the focus was on reading in English at Grade 5 level, as this is the language of instruction in nearly all deep rural schools after Grade 3. The team made use of two separate oral reading fluency tests and two passages to test comprehension. In monograde schools, the entire class was given a written test, on the basis of which top, middle and bottom groups of learners were selected for a one-on-one oral reading fluency test. In total, 1790 learners completed the first oral reading fluency test, and of these, 878 learners who did sufficiently well on the first test completed the second, more difficult test. The research team made use of passages from approved Grade 4 and 5 textbooks with the first test of fluency passage deemed suitable for a learner at the end of Grade 3. When the test information was analysed, the researchers found that the Grade 5 learners could correctly answer about 4 of the 20 comprehension questions (20.7 per cent) with only 6 per cent of the sample of learners correctly answering more than 10 questions correctly. To provide insight into the kind of comprehension Grade 5 learners can do, they

differentiated the 20 questions into questions that required retrieval of explicitly stated information, questions that required making straightforward inferences from the passage, questions that required learners to integrate ideas and information across the texts, and questions that involved examining and evaluating the text. Only 32 per cent of learners demonstrated that they could retrieve basic information and only 26 per cent had developed the reading skills that would allow them to make straightforward inferences.

The second component of the assessment of the English reading of rural Grade 5 learners involved two tests of oral reading fluency. The rationale for testing oral reading fluency is that it has been found to be a good indicator of overall reading competence and that learners who read less than 50 words per minute struggle to make meaning of the texts they are reading.

Of the sample of 1790 rural Grade 5 learners who took the oral reading fluency test, 10 per cent could not read at all and did not seem to understand anything that the evaluator asked them in English. This finding is very significant as other studies such as the PIRLS had a floor effect and did not provide insights into the variability among so-called poor or non-readers. A further 11 per cent of learners could read only a few words and at a very slow pace. The NEEDU report inferred from this that 'nearly 22 per cent of Grade 5 learners tested are illiterate'. On the other hand, 49 per cent of learners tested were close to the appropriate level of fluency, even if their comprehension remained limited.

Building on the 2013 NEEDU research on oral reading fluency, researchers have deepened and extended the analysis of reading underachievement in South Africa. Using oral reading benchmarks from the United States, Spaull (2015) compared the cohort that was tested in 2013 to primary-school children from one county in Florida in the United States. He found that the kernel density distribution of the Florida second language learners in Grade 1 was most similar to the distribution of Grade 5 rural South African learners (save for the substantial number of South African schoolchildren who registered close to zero).

Spaull (2015) concludes that the international benchmarking exercise of South African Grade 5 reading fluency reveals that rural Grade 5 non-English speakers' oral reading fluency is very low. With more than 40 per cent of learners reading at less than 40 words per minute, it is unlikely that these schoolchildren are able to make much sense of what they are reading. The NEEDU Grade 5 sample of school learners was on average at the same reading fluency levels as the slower Grade 2 second language learners in Florida. That said, in a subsequent follow-up analysis to the international benchmarking study, Pretorius and Spaull (2016) recognised that South African reading norms for second language might be considerably different from those often used in international studies. Their research suggests that some South African children who read slower than the international benchmark of

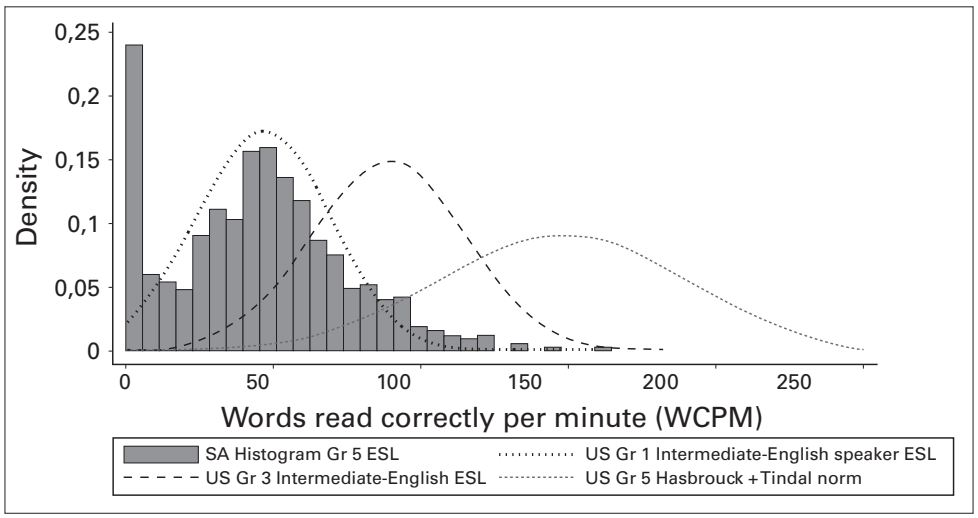


Figure 2.5: Distribution of ORF scores (WCPM) for South African First Additional Language relative to Broward County ESL learners
Source: Adapted from Spaull (2015)

90 words correctly per minute are able to understand what they are reading, and that the South African threshold for becoming a reader who can comprehend what he or she is reading is more likely to be around 79 words per minute.

This additional analysis, while providing insight about the usefulness of an international benchmark for oral reading fluency, nuances the picture but does not fundamentally alter the message. Not only are almost half of South Africa's children not able to understand what they are reading, but they lack an underlying skill, that is, fluent and accurate reading of words on the page.

Annual National Assessments

Following on the earlier field trials and the first wave of sample learner testing (the Systemic Evaluations), the Department of Basic Education embarked on an ambitious task of testing all South African children in public schools in Grades 1 to 6 in literacy and numeracy/mathematics. From 2011 to 2014, the Annual National Assessment (ANA) included two testing protocols: the universal testing of all public school learners administered and marked by teachers and a separate 'verification' testing process moderated by an external agency. The purpose of the randomly selected verification ANA was to validate the results that emerged from the universal tests, particularly to check the administrative process and the rigour of the marking undertaken by teachers in their own schools (Department of Basic Education, 2011).

Van der Berg (2015) observed that while problems related to the inter-grade comparability of ANA test scores limited the ANA's usefulness for measuring learning gains, in his view, these tests provided meaningful information on learning gaps between children in a single grade and in a single year.

For the Department of Basic Education, ANA would 'make a decisive contribution towards better learning in schools' by encouraging teachers to assess learners using appropriate standards and methods, encourage better targeting of school support by provinces and districts, recognise and publicise instances of successful performance and encourage parents to pay attention to learning outcomes.

The Department of Basic Education's *A Guideline for Interpretation and Use of the ANA Results* (2011) set the benchmark for performance — 'an acceptable level of performance is linked to a score of 50 per cent and above, since a score of 50 per cent and above is referred to as achieved. Learner performance in ANA is therefore reported in terms of the number of learners who obtained a score of 50 per cent and above, which includes performance labelled "achieved (50–69 per cent)" and "outstanding (70 per cent and above)"'.

Table 2.1: Percentage of learners at acceptable level of performance for grade and subject in 2011

Acceptable level of performance	
Grade 3 Literacy	31%
Grade 3 Numeracy	17 %
Grade 6 Language	15 %
Grade 6 Mathematics	12 %

Source: Adapted from Department of Basic Education (2011)

These not-unexpected results are a profound indicator of South African primary education. In the 2011 ANA, 70 per cent of South African Grade 3 schoolchildren were not at an acceptable level. Just over half of all Grade 3 learners scored below 35 per cent on the ANA test, and a further 16 per cent scored between 35 per cent and 49 per cent. Only 15 per cent of all Grade 6 learners scored above 50 per cent, the minimum acceptable level of performance.

What exactly does it mean that only 15 per cent of South African Grade 6 learners are at an acceptable level of performance in language? More precisely, what is it that 85 per cent of South African schoolchildren cannot do? The exemplar question (Figure 2.6) from the 2011 ANA test, published by the Department of Basic Education, provides some clues. The passage and related question required learners to have a modest English vocabulary and be familiar with the meaning, such as 'patient'. The learners would need to be able to read with sufficient fluency and understand the basic features of grammar to capture the meaning of the passage.

They would need to be able to read back and forward to extract information that was explicitly stated at the beginning of the passage. The one relatively minor area of ambiguity relates to the conjunction 'and' in the paragraph and the causal inference in the question. And while the teacher, learner and school names are English, the setting and the set of events discussed should be familiar to all children who have attended school for six years.

MARK WILSON'S SECRET

The Grade 6 learners at Springvale Primary School were excited because they had a new teacher. Her name was Miss Martin. She was patient and kind and the children loved her!

One day when Miss Martin was marking the register she noticed that every now and again Mark Wilson was absent. The class told her that Mark was "not good at school work."

Then, one cold day in June, Mark arrived at school with a big smile on his face. "May I read please, Miss?" Mark asked confidently. Mark's secret was out. He had finally learnt to read.

LEARNING OUTCOME 3: READING AND VIEWING

1. Draw a circle around the letter of the correct answer. (1)

The children loved Miss Martin because she …

A. greeted the class and read the story.
B. noticed that Mark was absent.
C. was patient and kind.
D. visited Mark's house.

Figure 2.6: Exemplar Question from the Annual National Assessment English Language, Grade 6
Source: Adapted from Department of Basic Education (2011)

The 2014 ANA aggregate results by province for Grade 3 and Grade 4 (Home Language and First Additional Language) provide additional information on the prevalence and patterns of underachievement. At a national level, the ANA results suggested that 43 per cent of Grade 3 learners failed to achieve a score at or above on the language test (in all languages). There was large province variability, with more than 70 per cent of Gauteng Grade 3 learners 'passing' compared to 42 per cent in Limpopo. Once again, the provincial differences are an indicator of the bimodal distribution, with provinces with the largest proportion of Quintile 4 and 5 schools showing a consistently larger proportion of both Grades 3 and 4 learners

achieving scores above 50 per cent. The other big story from the 2014 ANA is the huge drop-off between Grade 3 and Grade 4 in the language scores, just as most learners change from instruction in the Home Language to receiving instruction in English. Nationally, the discrepancy between the proportion meeting the minimum standard in Language in Grade 3 and the proportion meeting the standard in Grade 4 is large. By Grade 4, only one in three schoolchildren met the minimum academic standard in the language that was used for teaching and learning at school.

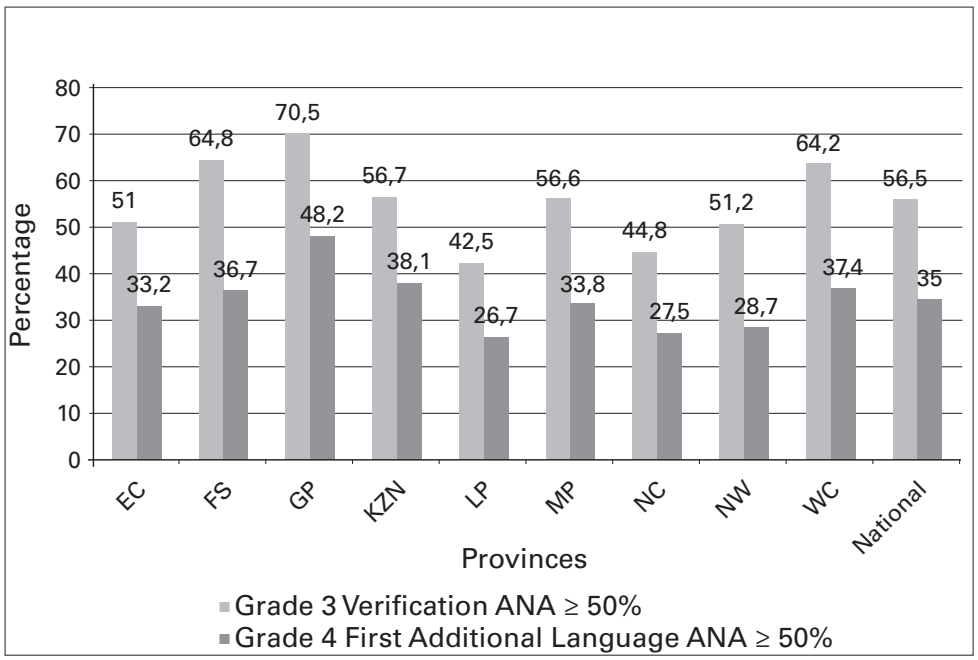


Figure 2.7: Percentage of Grade 3 and 4 that achieved 50 per cent or more on Language and First Additional Language Annual National Assessment by Province, 2014

Source: Data for this figure comes from Department of Basic Education (2014)

Reading Catch-Up Programme

Are the findings about South African schoolchildren's reading underachievement from PIRLS 2011, the NEEDU reports (including subsequent secondary analyses) and the Annual National Assessments consistent with other independent research findings? If we examine the evidence from the one independent source, the endline learner test results from the Reading Catch-Up Programme (RCUP) (Fleisch, Pather & Motilal, 2017), the answer would be yes. While not intended as a study to explore reading achievement, the RCUP study provides useful insights into the shift from Home Language to English as a First Additional Language, the questions related to the variability of reading skills among schoolchildren in Quintile 1–3 schools,

and the underlying reading subskills that learners have or have not acquired during their first three years of schooling.

The Reading Catch-Up Programme study was a randomised control trial (RCT) designed to test the impact of a system-wide remedial English reading programme for Grade 4s. The RCT was conducted in 100 Quintile 1–3 schools (40 treatment, 60 control) in the Pinetown district of the KwaZulu-Natal Department of Education. To measure learning gains, all Grade 4 learners completed pre- and post-tests that included spelling, comprehension and writing components. The test was benchmarked against the official curriculum, but included Grade 2, 3 and 4 English First Additional Language level questions, with a few at Grade 5 level (Fleisch, Taylor, Schöer & Mabogoane, 2017b).

The study showed that while schoolchildren in both the intervention schools and the control schools improved their performance between the pre- and post-tests, most learners in the district continued to perform substantially below curriculum expectations. This was most clearly evident in their performance on the 22-item spelling test (Figure 2.8).

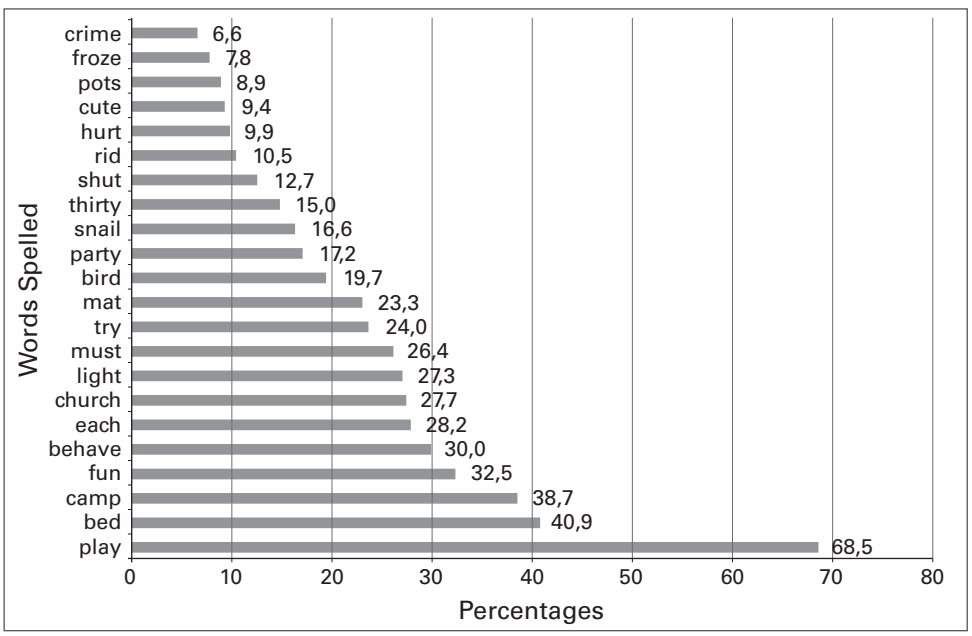


Figure 2.8: Scores for post-test: percentage of learners who spelled the word correctly
Source: Fleisch, Pather & Motilal (2017)

Only 40 per cent of Grade 4 learners could correctly spell the word *bed* and less than a third could spell the work *fun*. In terms of the official curriculum, these types of monosyllabic words with the simplest consonant-vowel-consonant (CVC) structure should be mastered by the middle of Grade 2. The difficulties with correct

spelling of simple English words strongly correlated with the schoolchildren's difficulties with comprehension. These findings are consistent with the challenges South African English second language schoolchildren have with reading in general, but particularly spelling (and reading) in English.

Table 2.2: Spelling errors by linguistic characteristics

Linguistic characteristic	Bed (per cent)	Fun (per cent)	Rid (per cent)	Mat (per cent)	Average (per cent)
Alphabetic	13 %	9 %	8 %	9 %	10 %
L1 interference	16 %	36 %	17 %	61 %	33 %
Vowel error	66 %	54 %	71 %	26 %	54 %

Source: Fleisch, Pather & Motilal (2017)

A secondary analysis of the types of spelling errors these Grade 4s made provides more information about the particular challenges that these children have in learning to read. A rough coding of linguistic features of the spelling errors (Table 2.2) for four three-letter words suggests that between 8 per cent and 10 per cent of learners could not match the first phoneme in the spelling word, for example [*b*], with the first letter of the word *bed*. These learners could be described as being at an alphabetic stage, with partly or semi-developed graphophonemic skills. Without the mastery of the basic grapheme-phoneme relationship these schoolchildren are unlikely to be readers in any meaningful sense. Figure 2.9 below illustrates the responses of these learners to the task to spell the words *bed*, *rid*, *fun* and *mat*.

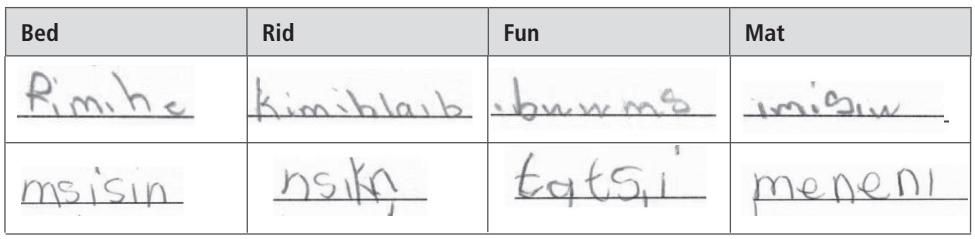


Figure 2.9: Alphabetical errors (pre-phonic)
Source: Fleisch, Pather & Motilal (2017)

The secondary analysis of the RCUP spelling data adds to the weight of evidence both of the pattern and prevalence of the non-reader group at the early stages of the Intermediate Phase of schooling. Like the work on oral reading fluency, the RCUP study provides insights into the unique linguistic challenges learners have as they move from learning to read in isiZulu and reading in English at the early stages of middle school.

Conclusion

What have we learned from a re-examination of the various sources on early grade learning, particularly reading outcomes? The evidence presented in this chapter serves to confirm the two widely accepted findings (Fleisch, 2008; Spaull, 2013). South African schoolchildren are performing well below their international counterparts and below our own curriculum benchmarked expectations. And the ubiquitous inequality problem, the bimodal pattern of achievement, closely tracks relative wealth and poverty.

But a careful analysis provides three new insights. First, that while achievement in Home Language learning is clearly below expectations, children's reading performance as they enter the Intermediate Phase of learning and shift to English is far worse in the new language of teaching and learning. The evidence shows that simply improving reading skills in the Home Language, while critical for improvement, would be insufficient. At a policy level, we need to be thinking about improved teaching in Home Language reading and improved teaching, even in the early grades, of English as the main First Additional Language.

Second, because of the 'floor effect' in many of the international tests and even our own large-scale testing exercises we have tended to assume that the performance of children in the poorly performing part of the distribution is more or less homogeneous. There is growing evidence that suggests that the pattern of reading performance in primary schools in township, rural and informal settlement areas should be further disaggregated. What is likely to emerge from this segmentation is a picture of a distinct subgroup of learners (of between 8 per cent and 10 per cent of learners) in these schools who are genuinely illiterate, having not gained the basics of reading, particularly the principles of the phonics system.

Third, the common assumption that the underlying problem in reading or literacy is about comprehension, and that children lacked 'meaning-making' skills but are able to 'bark at print' or decode, needs to be fundamentally reconsidered. Recent evidence suggests that a substantial proportion of our early grade readers are not able to 'bark at print', as most clearly evidenced by the high proportion of children with very slow oral reading fluency. The binary analysis that suggested that the basic decoding skills are in place but the comprehension skills are lacking does not stand up to the research evidence. Both appear to be challenges for young readers. The majority of South African schoolchildren are struggling with most, if not all, components of reading in their home and additional languages, that is, phonological awareness, vocabulary, phonics, fluency, comprehension and writing.

Chapter 3
Why policy fails, and the pivot to instruction

Introduction

How can we improve primary school achievement, and specifically the reading performance of the majority of children across the entire school system? This seemingly simple question is at the centre of this book. Seemingly, because when the question is unpacked, it emerges that the question is as complex, if not more so, than the question about why South Africa has such an extreme pattern of underachievement in primary school learning.

To answer the question, we actually need to answer three separate but interrelated questions. First, what are the predominant instructional practices associated with the teaching of reading (and mathematics) in primary schools? In *Primary Education in Crisis*, I attempted to answer it by exploring four aspects of classroom practice: teachers' knowledge; use of time in the classroom; the availability of resources; and teaching methods. In many respects, this proved to be an unsatisfactory account of instructional practice as it provided skimpy and disconnected insights into the systems of instruction.

The second question relates to 'what works' to raise the bar and narrow the gap of reading achievement. I did not even begin to address this question in the previous book, suggesting only that 'what works' cannot be inferred from educational theory or even international experiences. I suggested two reasons for this. Genuine knowledge for instructional practice must be developed through a combination of theory and empirical evidence, the one informing the other. I hold the view that the most rigorous empirical evidence for 'what works' cannot come from studying outliers, as is often the case in the school/teacher effectiveness research, but rather from experimental research, that is, randomised control trials. While international experiences are a good starting point, given the impact of contextual and historical realities, it is essential that policies or programmes designed to improve primary school reading and mathematics are demonstrably effective in the context they are designed for.

The third question relates to how best to translate the 'change knowledge' from the large-scale experiments into a system-wide implementation. This question recognises that doing this simply by virtue of the fact that we might have insights

into proven good practice from the South African context, would not be sufficient. Specialised implementation research is needed as we move to scale. This requires careful cost analysis and realistic assessments of system capacity, and the political will to reprioritise government spending.

The central argument of this book is that South Africa's priority for education change should be system-wide change of instructional practice in the early grades. By this I mean fundamentally altering the methods, routines and rhythms of classrooms, pacing, sequencing of activities and tasks, the kinds and quantities of educational materials that line the shelves and walls of these classrooms, teachers' skills and knowledge about the new practice, and more. The fixation in the media on improving Grade 12 examination results, particularly Matric pass rates and narrowing the gap between underachieving and achieving schools, important as they are, distracts policy-makers and system-managers from the critical business of planning and managing deep instructional change.

Primary Education in Crisis presented a picture of bimodal achievement patterns and those factors, both proximal and distal, that contribute to poor learning outcomes for disadvantaged learners. The term '**proximal**', from the Latin *proximus*, refers to those factors closely related to learning outcomes, such as appropriate instruction and time, and feedback on learning tasks that are similar to those that will be tested. The term '**distal**', from the Latin *distare* (to stand away from), refers to the factors that may have impeded children's test capability, that form barriers or hurdles, such as ill health or poor school management. While previous work gave equal weight to distal and proximal factors in trying to make sense of the huge discrepancies between middle-class and working-class children's academic achievement, this book shifts the debate towards a close examination of those instructional practices that are most immediately linked or are proximal to what determines how children learn. Again, what is proximal are the instruction practices that young children experience day in and day out in early grade classrooms. This is not my original insight, but draws heavily on the work of a group of American scholars, particularly David Cohen, Stephen Raudenbush, Deborah Ball, James Hiebert, Tom Hatch and Richard Elmore. While differing in emphasis, their work has persuaded me to shift my lens away from a preoccupation with the way in which home and community factors constrain achievement and towards the very real and direct relationship between instruction and underachievement.

This chapter is divided into three sections. Framed by the problem of underachievement in reading and mathematics, it first explores the changing policy environment as it has unfolded over the past two decades. The purpose of this policy analysis is to demonstrate that the continuing education crisis cannot be explained by the absence of attempts on the part of state actors to remedy the crisis. The fact that policy and programmatic interventions have largely failed to ameliorate the situation is explained, at least in part, by the state actors' use of incomplete change theories. The second section explores the limitations of change theories implicit in

two influential research traditions, one emerging from the field of education economics, the other from the school/teacher effectiveness research. Both traditions have had significant influence on both policy and programmatic responses. While these research traditions have furnished the education community with what is clearly the most robust evidence, the theories of change that underlie this research permit only *partial* insights into what is wrong and how to fix it. Building on the first and second sections, the third section of this chapter begins the conceptual development of the notions of instruction and systems of instruction, key concepts that inform the theory of change at the heart of this book.

Policy interventions

Absence of new policies is clearly not the reason for the continued underachievement in primary school reading and mathematics. South African schools have been subjected to wave after wave of well-intentioned, but poorly conceived and poorly implemented, half-used and prematurely abandoned education policies and private sector supported interventions. The first wave focused on teacher rightsizing and school governance reform. There have been three major curriculum reforms processes that began with Curriculum 2005 in 1997/8, followed by the National Curriculum Statement (NCS), and its re-contextualised version in the Curriculum and Assessment Policy Statement (CAPS). There have been efforts at improvement via labour relations agreements. These were intended to strengthen the incentive systems through job progression in the form of Occupation Specific Dispensations. Also aimed at teachers, the integrated teacher development policy made it mandatory for all teachers to acquire a minimum number of professional development points over a three-year cycle.

Some of these government policy initiatives have been designed to balance capacity-building with accountability. Accountability policies include the Integrated Quality Management System and more recently the National Education Evaluation and Development Unit. After more than two decades during which few classroom visits took place, there has been a renewed commitment to the role of supervision and even inspection; so, classrooms again come under external bureaucratic scrutiny.

Standardised learner assessment has also resurfaced in the policy arena. There has been increased attention to external testing, beginning with the provincial systemic evaluations, which were replaced with the universal testing in the Annual National Assessments from 2011 to 2014. Starting in 2001, the then Department of Education began releasing the Grade 3 and Grade 6 systemic evaluation results. While these government studies covered only a sample of schools, they sent a powerful signal to all schools about curriculum standards. A number of provinces have gone beyond the national assessment introducing their own more comprehensive assessment processes. The Western Cape, for example, had included all primary schools, rather than a sample, providing feedback to each institution on their

learners' relative performance compared to other schools in the district, region and the province as a whole. Despite serious teething problems, Annual National Assessments from 2011 to 2014 did play an important role both at a national level, as a reminder of the achievement challenge, and at the district and school level, as district managers and teachers grappled with the first real national benchmark of primary school achievement.

There have been extensive policy formulations and programmes for teacher development in particular. Possibly the most significant are the Funza Lushaka initial teacher education funding scheme, the funding of the Advanced Certificate in Education in the mid-2000s and the overarching frameworks of teacher development and the South African Council for Educators (SACE) teacher professional development point system. Funza, launched in 2007, was designed to promote initial teacher education for public schools through the provision of generous full-cost bursaries. Priority areas for these bursaries included Foundation Phase, African languages, English language and Mathematics. While the most ambitious of all the teacher development initiatives, it certainly was not alone. The substantial funding provided to universities for the Advanced Certificate in Education programmes in the 2000s, particularly those designed to upgrade teachers to teach literacy and mathematics to Foundation and Intermediate Phase students, dramatically increased the number of in-service teachers 'trained' to address the crisis in primary school achievement.

In addition to generic education policies and initiatives, the Department of Basic Education initiated a range of programmes that were specifically designed to address underachievement in primary schools. These have ranged from the National Reading Strategy that was launched in 2008 to the most recent government/private sector joint venture referred to as the Primary School Reading Improvement Programme (2017) (Table 3.1).

Table 3.1: National initiatives to address primary education underachievement

1. National Reading Strategy
2. Teaching Reading in the Early Grades: A Teacher's Handbook
3. Western Cape Literacy & Numeracy Strategy 2006–2016
4. Foundations for Learning Campaign
5. Early Grade Reading Assessment
6. Systematic Method for Reading Success
7. Drop Everything & Read Campaign
8. Certificate in Primary English Language Teaching
9. Primary School Reading Improvement Programme

Source: Adapted from Spaull et al (2016)

Provinces have adopted their own parallel strategies. In the mid-2000s, Limpopo, for instance, tried to retrain serving teachers from scratch, opting to send a substantial number of mid-career teachers to the Wits School of Education to complete a full-time four-year Bachelor of Education degree. Small teacher-development initiatives have also flourished around specific initiatives or programmes. The Western Cape has a series of programmes, including the Literacy and Numeracy Strategy (LITNUM)

This list of policies and programmes, all of which were or are intended to improve early grade learning, is not exhaustive. What are we to make of this extensive policy activism explicitly and directly related to improving primary school academic outcomes?

Given the scale, scope, level of, and increase in, funds allocated to the public school sector (Figure 3.1), it is clear that policy activism is more than 'symbolic' (the principle purpose of early curriculum policy, as Jansen (2002) has suggested).

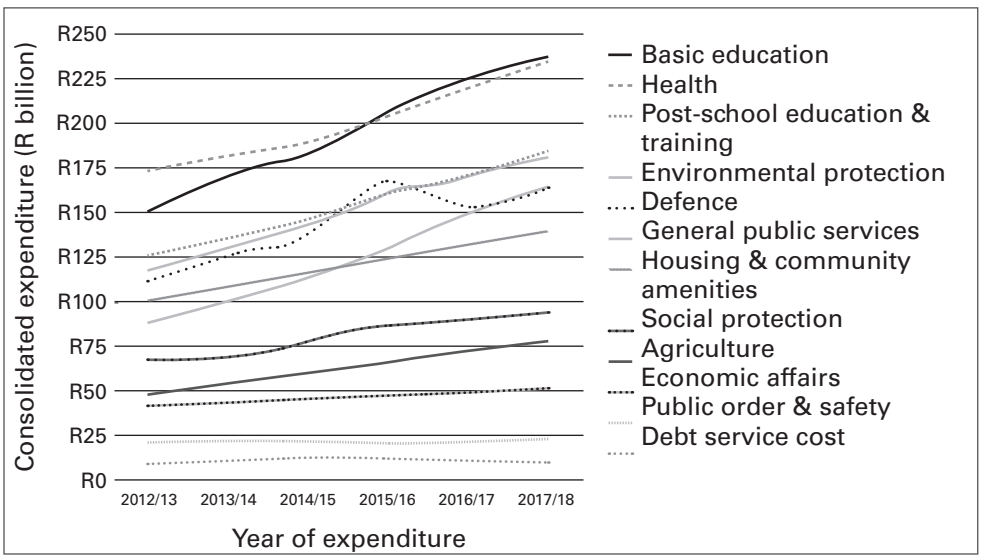


Figure 3.1: Government spending on education and other major services, 2012–2017
Source: Veriava (2017)

There are various other explanations for the failure to see improved learning outcomes at scale offered by a growing body of policy-researchers in South Africa. For example, researchers attribute the failure of the Integrated Quality Management System (IQMS), to the residual experiences that teachers had with the previous inspection system (Biputh & McKenna, 2010); technical weaknesses of the IQMS policies, such as the lack of co-operation, insufficient funding, and weak evaluation capacity at school level (Khosa, 2010); and financial disincentives (Turnbull, 2010).

It is important to recognise that policy and programme failure is not a uniquely South African problem. Conn's (2017) meta-analysis of different kinds of policy

and programme interventions and their impact on learning largely confirms this story.

> [T]his meta-analysis highlights the fact that many of the interventions under study work best as complementary interventions to improvements in pedagogical methods. For example, interventions in class size reduction, student tracking, or textbook provision presume that teachers know how to best take advantage of a smaller class size or a more homogenous classroom and that they can integrate these textbooks into everyday instruction. Furthermore, interventions that employ teacher-pay-for-performance schemes are partially predicated on the assumption that "newly motivated" teachers know how to improve the performance of their students. In addition, programs meant to increase learning through increased student and teacher attendance (eg conditional cash transfers or teacher financial incentives) may also be less successful if the quality of instruction itself is not changing. (Conn, 2017:888)

At least four kinds of additional explanations need to be explored.

1. The first focuses on the quality of the policies themselves. Are they well designed, conceptualised for the South African education contexts and have they been thoroughly field-tested prior to large-scale rollout? In most instances, time has not permitted extensive piloting of prototypes, let alone rigorous testing in randomised trials.
2. The second relates to implementation. Even if policies have a strong evidence base, failure comes through lack of political will or organisational capacity and know-how, particularly inadequate ongoing engagement with key stakeholders.
3. Third, policies fail because of policy fatigue and leadership churn. The 'recipients' of policy become disillusioned and adopt a wait-and-see attitude that says 'just as with the policy before it, this too will pass'.
4. Finally, policies fail the key test of alignment and coherence. This may be viewed as a subset of implementation failure and policy fatigue, but has a distinct character. New policies get absorbed into old embedded practices; this happens even more rapidly if there are overlapping and contradictory messages in the policy space.

All these explanations certainly have merit. However, for the purpose of this book I will offer an additional explanation: policies and intervention programmes often fail because they do not get at the instructional core.

Beyond input-output and school effectiveness models

Why are these policies and programmes unlikely to impact the instructional core and turn around academic underperformance? There are obviously a variety of reasons,

factors that South African researchers are beginning to discover. At least a part of the explanation for the failure is the flawed conceptual model that informed much of the education policy thinking.

Over the past 20 years, researchers, particularly those within the field of the economics of education, have assumed that the primary challenge of education was the identification of the 'key' input drivers, particularly those associated with unequal funding and resources. It was often assumed that once the basket of 'key' resources had been identified, the policy community could translate these into policies and implement them — resulting in improved achievement. Resources in the school setting are about funds or the things that funds can purchase, such as additional teachers to reduce class sizes, training for teachers, textbooks, tablets and school libraries, to mention a few that have been discussed in South Africa.

Possibly the most influential of these South African studies was undertaken by Anne Case and Angus Deaton (who subsequently won the Nobel Prize for Economics). Case and Deaton's (1999) study was a direct response to a raging international debate about the relationship between resources and learning outcomes, and particularly it responded to World Bank economist, Eric Hanushek, who is often associated with the claim that 'resources do not matter'. Using a variety of education and household datasets, particularly from the Department of Education and Training just before the 1994 election, their study 'differ[ed] sharply from what is often thought to be a consensus that school resources do not matter very much' (Case & Deaton, 1999:1080). Their findings showed that, controlling for class background, school inputs, particularly lower pupil-teacher ratios, were strongly correlated with improved learning outcomes. The implication: spend more and hire more teachers (of the same quality) to improve system-wide learning outcomes. Coming out of the apartheid system, this conclusion seems self-evident, and it continues to resonate with the recent discovery that over 10 per cent of all Foundation Phase classes are excessively large, with 50 learners or more (Spaull, 2016) per classroom. But nagging questions remain. Will just providing more teachers, teachers teaching in the same way, or one textbook per child, or tablets for all, really get children to read better or become more competent in the four basic numeracy operations?

While the input-outcome approach was used in some of the more influential early post-apartheid studies, economists began to recognise that resource differences alone did not tell the whole story. Luis Crouch (a senior consultant from the United States) and Thabo Mabogoane (then a locally based economist) (1998) found early on that in-school 'process' factors — things that economists normally struggle to measure — played an important part in explaining why some schools serving working-class communities did better than others. They described these as the 'residual' that is left over when all the quantifiable variables have been taken into account in the production-function models.

The second generation of school-effectiveness researchers went beyond the large measurable 'inputs' such as learner-teacher ratios and tablets in classrooms, and began to uncover aspects of improvement within the 'black-box' of schools, what Crouch and Mabogoane (1998) called the 'residuals'. Taylor, Van der Berg and Mabogoane (2013) produced possibly one of the best examples of the South African version of school-effectiveness research. In their multi-year panel study, the researchers were able to discover a range of key in-school variables that explain why some schools serving poor learners are able to perform better than others, with the same level of resources.

The National School Effectiveness Study (NSES) found that effective schools serving poorer learners had the following features in common:

> An organised learning environment signified by curriculum planning for the full year; a functional timetable; good quality inventories for LTSM; low teacher absenteeism and up-to-date assessment records. (Taylor, 2011:1-56)

The newly discovered variables include aspects of curriculum management, such as curriculum coverage and curriculum pacing. The study also identified the frequency with which teachers made use of certain kinds of learning activities or tasks, for example, extended reading and writing. NSES also discovered key differences in teachers' content knowledge, between performing and underperforming schools, as well as school management variables such as the existence of learner support materials inventories.

The problem with the school effectiveness research, which Taylor and colleagues acknowledge, is that it could be misinterpreted by policy-makers as uncovering a series of policy levers. For example, if the research identified one of the statistically significant variables as curriculum pacing, that is, that the underperforming classrooms work at a slower pace with less content covered at a lower level of difficulty, it could be assumed that the policy lever would be to put in place a set of 'pace setters' that ensure that all classrooms are at the right 'pace'. This analysis is obviously too simplistic. The variable 'slow pace' is a visible manifestation of a set of internal and often less visible mechanisms. While slow pace is associated with low achievement, the real causal factor may be linked to teachers' responses to learners' actual cumulative underachievement in prior grades. If learners have not mastered the basics of the four arithmetic operations, or have very slow oral reading speeds, the observed pattern of slow pace serves to accommodate learners' real reading capabilities, even if it exacerbates the achievement gap.

This insight follows the argument that University of Chicago economist Stephen Raudenbush (2008) made about the production-function education policy research in the United States. Despite 40 years of extensive research using input-outcome (Model A in Figure 3.2), education economists have been completely unsuccessful in translating the 'variables', for example, textbooks provision, into policy interventions

that would lift learning outcomes system-wide. The problem, as Raudenbush saw it, was that the basic input-outcome model is incomplete. Rather than an input (e.g. textbooks or smaller classes) improving learning outcomes or achievement, it is actually improved instruction that generates higher learning outcomes. In Raudenbush's view, resources are certainly important, even critical, but they do not in and of themselves drive improved learning. They have to be used consistently to be part of the improvement cycle. Resources can be used well and they can be used badly. In the former case, they are productive; in the latter they make no difference. Resources have a role as a moderating or intermediate factor. In other words, it is teaching and learning or instruction that is the proximal cause of improved learning (Model B).

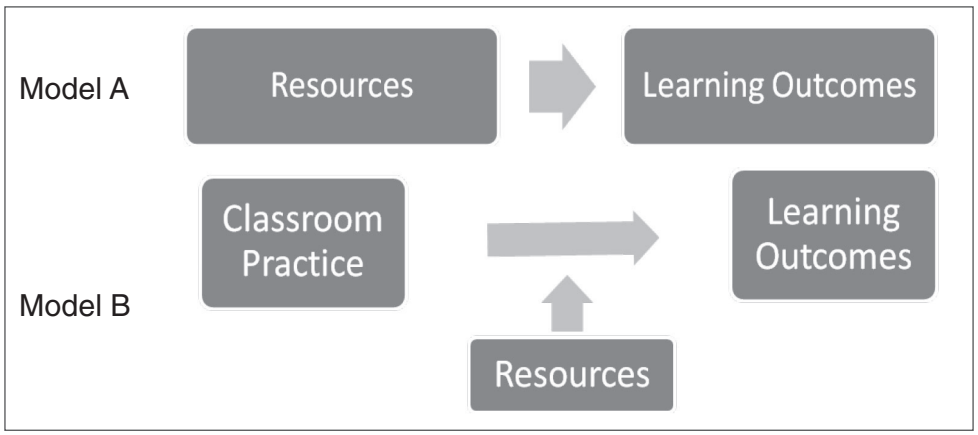


Figure 3.2: Models of education improvement
Source: Adapted from Raudenbush (2008)

While much of the concern in policy analyses (Case & Deaton, 1999; Fiske & Ladd, 2004; Crouch & Mabogoane, 2001; Van der Berg, 2009; Gustafsson & Patel, 2006; Taylor, Van der Berg & Mabogoane, 2013) has been around resources and more recently management indicators and their impact on outcomes, the model that begins with the relationship between classroom practice and learning outcomes and then views resources as a moderating variable pivots us to a focus on what Elmore (City et al, 2009) calls the 'instructional core'.

Instruction

But what exactly is 'instruction'? How does it build on and extend the variables discovered by school- and teacher-effectiveness researchers? Herbst and Chazan (2012) suggest that a theory of instruction has been around at least since the 1960s and that there is a strongly developed tradition in Europe, particularly in France,

called the study of didactique. It was only in the late 1990s that the concept began to gain popularity in the English-speaking research community. Drawing on the work of a range of scholars with regard to this and related concepts, this section does the work of unpacking the triadic concept of instruction.

The starting point is David Cohen, Stephen Raudenbush and Deborah Ball's (2003) frequently cited definition:

> Instruction consists of interactions among teachers and students around content, in environments. The interactions occur in distance learning, small groups in classrooms, informal groups, tutorials, and large lectures. "Interaction" refers to no particular form of discourse but to teachers' and students' connected work, extending through, days, weeks, and months. Instruction evolves as tasks develop and lead to others, as students' engagement and understanding waxes and wanes, and organization changes … Instruction is a stream, not an event, and it flows in and draws on environments—including other teachers and students, school leaders, parents, professions, local districts, state agencies, and test and text publishers. (Cohen et al, 2003:122)

It is worth unpacking each part of these characteristics of instruction with an eye to Lawrence Cremin's definition of education (Cremin, 1976). Instruction can occur in a wide range of institutional settings and configurations, but for our purpose the school classroom is our point of focus. That said, there can be little doubt that home-based parent-learner-content interactions play a critical role in consolidating school-based learning, particularly in the early grades and specifically in the process of learning to read. Instruction is clearly different from 'teaching' as it involves both what the teacher can do and actually does and what the learners actually do. It is about the specific kind of interaction between teachers, learners and educational materials. It is first and foremost a deliberate and sustained (but not always systematic) set of interactions that occurs around 'content', whether it be knowledge, skills, attitudes, values or even sensibilities. These intentional interactions are divided up into lessons or even smaller units as activities or tasks. Instruction, however, is not the same thing as lessons, activities or tasks as it can and usually does occur across time and between lessons, activities and tasks.

What is powerful about this formulation is that it recognises that instruction is about interplay or interaction. Not only does it move us away from learners and the limitation placed on them by their socio-economic circumstances, it moves us beyond crude assumptions about the impact of resources on outcomes, and it helps us see beyond the limitation of teachers' knowledge and motivation. Instruction is about the content to be taught, teachers' knowledge and skills (and attitudes) specifically regarding how and what to teach, the resources and materials that learners can interact with, and the amount and the kinds of learning activities or tasks that

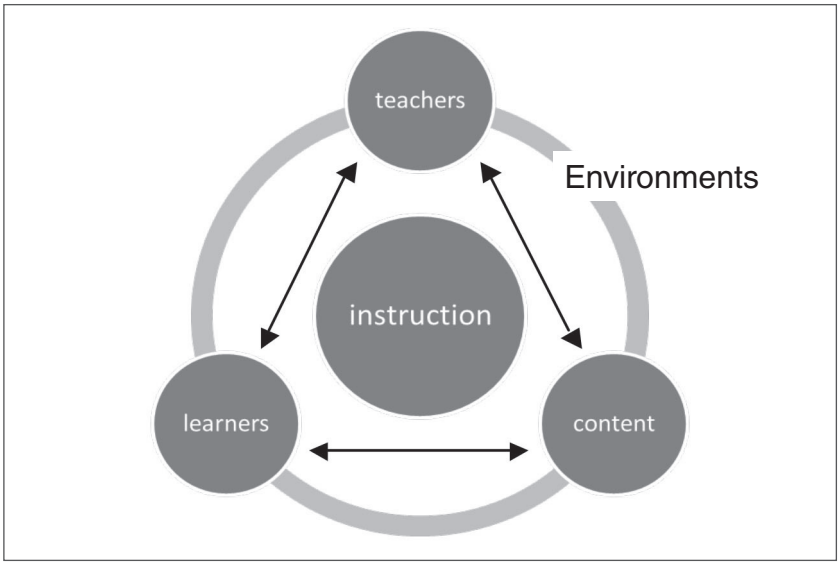


Figure 3.3: Model of instruction

learners are routinely required to do. To study instruction, we need to study the interconnectedness between components of the moving parts, each of which impacts directly on the other.

To illustrate the need to see instruction as interactional systems around content, it is worth examining a case. One of Ursula Hoadley's (2017) recent cases on Foundation Phase teaching is illustrative. The classroom observation came from a wider study of the best early grade teaching practices they found as part of the School Performing Above Demographic Expectation (SPADE) study.

> Another example comes from Teacher P, who in a small graded reading group, reads a story with a clear moral message about a boy who was too proud. The teacher begins by showing learners the back of the book, and asks what the book is about. Learners don't respond so she shows them the front cover and asks what the boy on the cover is doing. The teacher asks a series of cloze questions about the cover picture and title to which the learners provide single word answers in a chant. They chant the name of the author after her. They chant the title after her a number of times. The teacher begins to read. After a sentence she points to the speech bubble on the first page and asks "But class, I am looking here on the first page. What are these? These things?". The intercom interrupts with an announcement. She says "speech bubble" and the learners repeat after her. The teacher writes 'Question mark' on the board, explaining that when you ask a question you use a question mark. She also writes 'full stop' and 'exclamation

> mark' on the board and briefly explains what these look like. She writes 'comma', but doesn't explain this. She starts to read again.
>
> At the end of the first page the teacher stops reading and shows the learners how to turn the page. She shows incorrect ways of holding the book. She reads a sentence and then says that the learners must have respect for the book because "books make a person clever". Learners repeat after her "Books make a person . . . clever". She reads a few more sentences and then asks learners where they can get "a book for free to learn". Learners reply in chant "library". She shows learners again how to turn the page of the book. The teacher continues reading, stopping at times to explain words and at other times to discipline the rest of the class. The teacher continues reading and some learners join her in chorus. She stops to ask them what comes after the word "sjoe!" in the text, and they answer "exclamation mark". An individual learner reads while the rest of the learners follow in their books. The teacher goes to see what the rest of the class is doing while the learner reads. The teacher returns to the mat and takes up the reading and then the learners join in a chant. The teacher explains a quotation mark. She shows them a sentence in the story that is in quotation marks. The learners repeat the sentence after the teacher. They then continue reading. At the end the teacher gives the learners a comprehension exercise to do when they return to their desks. The exercise is unrelated to the reading that has been completed.

For our purposes, the case illustrates key elements of instruction as the unit of both analysis and ultimately of change.[1] There is now wide agreement that the purpose of group-guided reading methodology is to provide learners with opportunities to develop fluency, develop word-attack skills and read for understanding. The teacher's primary role when she uses this methodology is to assign learners into the small groups based on her assessment of the learners' reading abilities. All children with the same ability are placed in the same group. This is to ensure that the learners are reading texts at the right level. When the group-guided reading is actually happening, the teacher is expected to listen to each learner read, ensure that he or she is at the correct level, help with word-attack strategies for unfamiliar words and monitor progress.

In the case of Teacher P, although the learners had access to the critical resource, that is, the graded reader, it was of little use, as the case suggests that it may have been at the wrong level or the learners were unfamiliar with how to respond

1 Hoadley, using and extending Bernstein's analytic tools, argues that this and the other cases of the best-performing teachers suggest that teachers are at best conforming to the form rather than the substance of the ambitious national curriculum.

to the specific kind of text. The chanted responses to the cloze questions suggest that both the teacher and the learners were not familiar with the expected behaviours at the start of a group-guided reading exercise. The teacher's behaviour displayed lack of both knowledge and skills for the instructional activity. This suggests that Teacher P had probably not observed colleagues or read or heard much that would guide the use of the practice. It may be that the teacher had limited knowledge of the wider field of reading, including understanding the components of reading teaching (PA, phonics, vocabulary, fluency and comprehension), reading assessment, and classroom management — all knowledge domains that are necessary to make good use of the expensive resources. This knowledge and these skills are not sufficient; teachers like Teacher P need to be able to translate this knowledge into a set of activities or tasks that build on each other both between lessons and between tasks within the lessons. The intercom interruption stands in for all the many institutional and environment constraints that come to influence one or other component of the instructional triangle.

To identify the instructional core is not the same thing as understanding how it takes hold. The teaching methods, routines or rhythms of classroom instructional practices are constituted in two ways: by the forces that shape each component, and how components interact around content in the learning situation. A simple example illustrates this point. If a large proportion of schoolchildren in an overcrowded classroom are suffering from middle-ear infection that limits their ability to hear a teacher's directions, this will have a substantial effect on the enacted activities or tasks that the learners work on. It is likely to slow the pace at which learners begin the new task and it may lead to misunderstanding of the content that learners need to master. Ill health, whether chronic or acute, is but one among many factors that learners bring with them to the main moments of instruction. But the learners' personal circumstances, family and community characteristics do not necessarily predetermine the outcome of the instructional dynamic.

Teachers likewise come to the classroom with knowledge, skills and dispositions. They also come with their own personal histories and ideology. Their own schooling experiences often play a decisive role in shaping their instructional approach, possibly as much as their initial teacher education and subsequent learning in various communities of practice in the workplace. Although each teacher's personal history and work experience is unique, the common institutional spaces that they have occupied through their lives, first as learners and later as student-teachers, shape the collective nature of the instructional practice. Within the institutional spaces, there are structured and implicit incentives and sanctions that shape the common practices. There is also ideological 'baggage' that teachers bring to the set of tasks that make up classroom instructional practices.

Physical and educational resources and materials play a key role in the mix. These resources include the basics, such as physical buildings, the tables and chairs,

blackboards, textbooks, reading books, wall charts and computers, but they also involve the average number of learners in the classroom. Resources also include a whole range of other less visible but equally important assets, such as curriculum policy and other forms of external guidance, both in written and verbal form, from school managers and more distant departments. These resources are moderating variables, that is, they do not directly produce or determine instructional practices. However, they set the condition of possibility or place limits on the instructional practices. And as the provision of resources in public institutions depends heavily on standardised policies and prescribed official lists, their impact is to influence common practices across schools in working-class and poor communities.

For example, teachers' teaching tasks related to sound-letter relationships as one of the early building blocks of decoding may be linked directly to the teachers' beliefs about the capabilities of the learners. It was a commonly held belief that teachers should teach only one sound-letter combination at a time and that this needed to be done slowly so as to ensure that every learner had mastered this foundation knowledge. These beliefs may be re-enforced by low levels of mastery linked to low demands required in earlier grades. Teachers do not select educational materials that require extended reading because they believe and observe that most learners are not sufficiently fluent to manage the more challenging vocabulary in those texts. The absence of graded readers for appropriate level individual reading ensures that few learners become fluent readers. If teachers have low learning expectations, then missing one day a month is not going to fundamentally alter the aggregate level of learning in the class. Each component, whether it is a belief about what learners are capable of doing or the learning resources that are available, fits seamlessly into the instructional practice.

The instructional practice logic is held together by low expectation from parents, teachers and principals, justified by reference to children's low socio-economic circumstances and child-centred logic about teaching to meet the needs of learners. Once the power of the interlocking, mutual reinforcing system is properly understood, it becomes evident why single-input interventions are unlikely to shift instructional practice. An excellent example of this is the distribution of 'big books'. These innovative resources, which have been provided to schools by NGOs and are now a standard feature of the educational materials available to schools, seldom trigger a shift towards effective use of the technique of shared reading. More often than not these new resources get absorbed into the practice of whole-class choral reading. As a curriculum intervention, requiring teachers to cover the curriculum as measured by the level of completion of the DBE, workbooks trigger another form of 'mutual adaptation' (Berman & McLaughlin, 1974), children copying answers from the board into the workbook worksheets. The workbooks are now 'complete', but primarily as a result of the whole class copying from the board.

Systems of instruction

Responding to what they see as an American preoccupation with the deficits of teachers, Hiebert and Stigler (2017) have attempted to shift the debate by arguing that the problem in American education is not teach*ers*, but teach*ing*. Using Walter Deming's management theory, they argue that just like workers who were blamed for low productivity in the United States in the 1970s when it was the entire production process that was to blame, so too, the tendency is to place the blame on teachers and their deficits, either their inadequate knowledge or their lack of motivation. For Deming, the answer lay not in the hands or heads of workers, but in innovation in the business processes and the systems of production. The problems were as much with the way in which factory managers managed and how the production process was designed, as it was with the failure of individual workers on the factory floor. For Hiebert and Stigler, change involves all the ingredients of the system around instruction.

The theory of instruction outlined above points to the interaction of how teachers mobilise their knowledge and skills in the classroom, the educational materials available for learning, and the nature of the activities and tasks that learners do. Tom Hatch (2013) points us to wider dimensions of the systems of instruction that maintain existing instructional practices and are needed for improvement and innovation. He argues that the generic notion of capacity needs to be disaggregated into three distinct technical, human and social categories. He associates money and resources (including the class size) with the technical; the human with the distribution of skills, knowledge and dispositions of the key professionals in the systems; and the social with the relationships, levels of trust and collective commitment. In his analysis, simply having the resources does not mean these will be used well. Maximising their value depends on the abilities of people involved and the social connections between teachers, and teachers and other key stakeholders. While existing theories of change have focused on technical resources and more recently teachers' knowledge and skills (and motivation), the social relationship aspect of educational change has often been neglected. Strong working relationships, high levels of trust and a common understanding of a collective project of change are as important for Hatch as having the resources and upgrading teachers' knowledge. While the instructional practice remains at the core of change, social capital is needed not only in the classroom, but in the school, the district and at the head-office levels.

> It is extremely difficult to support 'innovation at the core' however, and to improve classroom practice throughout a school, a district, or a network of schools without organizational capacity. In turn organizational capacity depends on the distribution of resources across classrooms, the collective skills, knowledge

> and dispositions of teachers, staff and students, and the social networks amongst individuals and groups in the organization. Unfortunately, many schools and districts that lack the instructional capacity also lack organizational capacity. They lack the capacity to make improvements in the classrooms, and they also lack the capacity to make significant changes in their organizational structures and practice. In other words, it takes organizational capacity to build instructional capacity. (Hatch, 2013:36)

In Hatch's view, in addition to requiring new organisational capacity at various levels to make instructional change possible, dense and interconnected social networks are a fundamental ingredient.

Instruction in systems

Cohen and colleagues (2003) point out that 'much instruction that researchers had associated with individual teachers' work have collective features, it was shaped by teachers' work together, by leadership, and by organizations and cultures in which students and teachers worked', (Cohen et al, 2003:122). While these practices may appear idiosyncratic, closely associated with each individual teacher's unique personality or style with each group of learners, there are clear patterns and common features that cut across groups of teachers, even national cohorts of teachers. In the main, these patterns and features are relatively stable, but do modify over time. While no two teachers' classroom practices are identical, similar teacher training experiences, standard sets of classroom resources, and commonly held beliefs about learners and about the national curriculum give instruction a system character. There is no better example of this than in Carnoy's (2007) study of teaching in Latin America. In Cuba, Chile and Brazil, Carnoy and his colleagues found that each country (and subsystems within these countries) had distinct instructional practices most clearly manifested in discrete patterns of seatwork, group recitation, individual recitation, individual and group work, whole-class activities and time for transitions and interruptions. These patterns, in turn, can only partly be explained by constraints such as class size and the availability of textbooks and other learning resources. Bruce Fuller and colleagues (1994), after working in Botswanan junior secondary schools in the late 1980s and early 1990s, observed that institutionalised pedagogical routines were reinforced by an assemblage of outside pressures.

In certain specific circumstances, 'innovations' can rupture prevalent system-wide instructional practices. But more frequently, external interventions, such as input policies, designed to shift instructional practice have little success, save for superficial adjustments that allow teachers to appear to be co-operating. Smaller class sizes or new book series, on their own, often change little in the instructional practices; teachers seldom use the small classes as an opportunity to rethink

instructional tasks. Books that are too demanding, beyond the content knowledge of the teacher, are simply left in the storerooms or brought out on the day of inspection.

Conclusion

This chapter started out with the question, why policy, with all its potential to drive improvement of learning, has largely failed. To say that it has failed to improve learning, however, is not to say that it has not succeeded in other ways. There is a lot of evidence to suggest, and my experience has shown, that South African schools are substantially better institutions than they were at the eve of the democratic elections in 1994. While 'mud schools' still exist, the vast majority of schools are in good condition, with formal buildings, electricity and running water, and staffed with qualified, often well-intentioned teachers. Millions of children receive meals at school each day during the school year. Notwithstanding cases to the contrary, many schools are relatively safe spaces for many young people in unpredictable, often unstable and resource-limited communities.

Policy, however, has failed to substantially improve learning outcomes. The turn to the concept of 'instruction' allows us to see that it is what happens in the intentional activities and tasks performed inside the classroom that has the greatest potential to improve learning. In the triadic concept of instruction, this is the deliberate, sustained (and systematic) interaction of teachers with learners around content. The smallest units of these interactions are the instructional task or activity situated within the lesson. Therefore, instruction is about teachers bringing their knowledge (general and subject specific), their understanding and skills in pedagogy, their dispositions and sensibilities, together with sets of appropriate resources or educational materials, to the interaction, and thus to the tasks that enable learners to do the learning. As with teachers, it is critical to understand what children bring to the activities and task, as both influence how instruction impacts learning.

The triadic theory of instruction does a number of key things. It enables us to recognise that none of the components of instruction alone, that is, the provision of educational materials or upgrading teachers' knowledge or skills, could lead to real improvement. As Richard Elmore reminds us, upgrading one of the components without upgrading the others is unlikely to have much impact. The focus on instruction as the core work of schooling enables us to differentiate between approaches to improvement that are likely to yield improvement (those that are proximal to instruction) from those that are shiny and new but are at a distance from instruction, which will touch only lightly on the core business of teaching and learning. These include interventions such as school leadership training and building computer centres. This formulation of instruction points directly to the need to understand the move from surface to substantive, to use Hoadley's (2017) phrase and, by extension, rigorous research to develop 'proven' instructional programmes

that have taken into account all components of the instructional process for the South African schooling environment. The big idea around change at the instructional core is the need for externally developed, evidence-based programmes that serve as exemplars of substantially different kinds of instructional practice. These programmes, if embraced and internalised by teachers, over time will lead to sustained improved classroom practice. And this, in the theory of instruction, is the precondition for changing the patterns of learning underachievement.

The triadic theory of instruction certainly has limitations. It does not provide the analytic tools to go deep into the instruction process itself. It cannot, for example, really help to unpack the nuances around pacing, sequencing, progression, selection, coverage and a host of other challenges, specifically around the relative cost-effectiveness of specific core methodologies. For that we must turn to the theoretical advancement supported by empirical studies undertaken by colleagues such as Hoadley (2017) and her work on evaluative rules and regulating criteria; Venkat's (2013) research on temporal range; and Shalem and Slonimsky's (2010) work on ordering logics. That is not the purpose of this chapter or this book, however, and to recognise the importance of this conceptual and empirical research on other aspects of instruction is not to discount the value of articulating the more general concepts.

Chapter 4
Practices of literacy teaching

Introduction

In order to make sensible headway with the problem of how to improve learning performance in primary schools, we have turned to the problem of teaching, or what Richard Elmore has called the 'instructional core'. The assumption is that what really makes a difference to achievement levels are those factors or forces that are proximal to learning, that is, what teachers do, or specifically those things they do that are directly related to the teaching and learning of the material and/or skills to be learned. But before we can begin the journey on how to improve the 'instructional core' we need to understand precisely what it currently is.

This would seem to be a rather trivial task, given that there are tens of thousands of classrooms doing it every day and that it would seem self-evident that we 'know' what the prevailing 'instructional core' actually is. Unfortunately this assumption is unfounded. We cannot infer what the prevailing instructional core is from the official curriculum and official policy documents. While government documents do provide us with insights into what education officials want teaching and learning practices to look like, generations of education researchers and simple common sense tell us that there is a huge gap between what policy 'says' and what practice 'does' at the classroom level. Equally, we cannot simply turn to the research literature. While the common phenomenon of classroom teaching and the instructional core occurs almost everywhere and all the time, few systematic studies have been done that simply document common or standard practice. While there is a large amount of research in the education field, surprisingly little has been published that systematically documents and analyses the ordinary, common instructional practices of typical South African teachers in their classrooms. Published research studies have tended to focus on the exceptional, the effective, the exemplary classrooms, teacher or practice. Our university-based researchers unfortunately tend to highlight practice that is cutting-edge, avant-garde, or against the grain.

Two instructional practices in literacy teaching[2]

In this chapter I present two distinctive literacy teaching practices. As with any attempt to collapse the huge complex terrain into a simple set of categories, the variety and unevenness of the actual world is lost. That problem notwithstanding, for the majority of children, possibly as many as 70 per cent, literacy or reading teaching in the first years of formal schooling is a long sequence of unsystematic lessons beginning with learning the names of the letters and basic letter-sound relationships, followed by endless lessons involving choral reading aloud and language worksheets. By the end of primary school, after sitting through hundreds of lessons, most children have learned the principles of the reading and writing system in two languages, but can identify only very basic information in short texts and write simple words and sentences.

The second literacy teaching practice is closely associated with what happens in many former white and Indian primary schools, and to a lesser extent in former House of Representative schools. Built on reading methods or techniques introduced in these schools in the 1960s and 1970s, this instructional practice focuses on getting the majority of children in the Foundation Phase to *read fluently,* but not necessarily to become critical or even to simply read for pleasure.

What is clear, however, is that while there is a strong relationship between the instructional practice that the school uses and the kind of reading children take up, the relationship between reading teaching and the eventual reading practice is not always determined. Schools that have very strong reading teaching practices struggle to help some children to become readers. On the other hand, there is a small cluster of children in schools which teach using chorusing, who despite their inadequate schooling become sophisticated 'readers' either because of their involvement with informal literacy networks in the home or community or because they have forged special relationships with teachers who facilitate their entry into the world of reading and writing. While some young people transcend or fail to take on the literacy practices of their schools and homes, they are the exceptions, the outliers. The purpose of this chapter, however, is not to focus on these children but rather to explore the typical or common literacy practices and how they link to patterns of reading achievement.

2 Kate Perry, using Brian Street's notion of literacies as social practices, describes three distinct kinds of literacy in Africa. The first is *mother-tongue literacy*, which in her view never really takes hold due to the absence of books in the mother tongue and the appeal of English. The second is *schooled literacy*, which is a rudimentary form of literacy in English or another European language, but in the absence of books in English is little more than the knowledge gained from choral recitation of text written on a blackboard. The third is *literacy practices beyond the school.* This practice Perry associates with the widespread use of appropriate libraries.

Choral reading aloud as the instructional core

Since Carol Macdonald's ground-breaking research on language and reading in the late 1980s, a not well-known group of researchers in South Africa has produced a body of empirical research that collectively portrays a stark picture of the typical or common pattern of reading teaching practice in disadvantaged schools. The large-sample standardised classroom research and small qualitative observation studies together present a consistent picture of the teaching methods and reading resources used in township and rural classrooms across the country. But more than this, the body of scholarship provides insights into why this practice emerged and the subtle shifts that have taken place even though the standard practice has remained remarkably stable.

While hardly ever referenced in this body of scholarship, possibly the most compelling study of reading teaching practice was commissioned by the Limpopo Department of Education. When Dr Aaron Motsoaledi was the MEC of Education in Limpopo in the mid-2000s, he became increasingly concerned about the level of reading proficiency in rural schools in his province. As part of a large school improvement project, he commissioned the Human Sciences Research Council (HSRC) to investigate classroom practice and the causes of poor reading performance. In response to the brief, the HSRC team (Reeves et al, 2008) designed a multifaceted study of reading practice including a large sample study of Limpopo primary school classrooms. The team of researchers did intensive and in-depth class observations in 160 classrooms in 20 schools and interviewed hundreds of teachers, principals, school governors, district officials and local university lecturers.

The study provides a valuable starting point for understanding the typical practice of the teaching of reading in disadvantaged schools in South Africa. The sample of schools the HSRC team visited was similar to rural and peri-urban schools in other provinces. Sixteen of the schools were located in deep rural areas where over half of the children reported living in huts or shacks. The schools varied in size from very small village schools with fewer than 170 children to large 1 000-learner schools in urban centres. Most of the schools the HSRC team visited had Grade R classes and all but one had functional school-feeding programmes. While the teacher-learner rates were reasonable at around 1:35 (going up to 44), the actual class sizes, particularly in the Foundation Phase classrooms, were high.[3]

Despite these challenges, the researchers were surprised to find that most teachers were very positive about their schools. Teachers were upbeat and more than willing to have the researchers observe their classes. On teachers' desks, the

3 Of the 80 teachers who were interviewed, about a quarter reported teaching classes with more than 50 learners. Despite a small number of very large classes and classroom shortages, the researchers found most of the schools were clean and neat and in basically good condition.

researchers almost always found the ubiquitous mark-books and in many cases copies of official government curriculum documents. They observed teachers teaching in a language that most children in the classes understood and could speak fluently, even if it was not always the children's home language. The researchers got the sense that the teachers they interviewed really tried to get parents involved in their children's academic progress.

One of the causes of Limpopo childrens' underperformance has to do with the language practices in the schools. While most schools in the province had returned to using the home language as the medium of instruction in the Foundation Phase, the researchers visited very few classrooms where English was being taught. As a result, very few Limpopo schoolchildren had acquired much English vocabulary, and most were not ready to use English as the language of instruction from the beginning of Grade 4. Only a few schools taught English words in the numeracy and life skills lessons. Visiting Intermediate Phase classrooms, the researchers observed that most teachers continue to use African languages in all subjects, either as the primary language or as part of extensive code-switching.

Teachers showed the researchers the home-language learning support materials that they used to teach reading. Teachers had a variety of teaching materials, mostly single copies, with only a third of the classes visited having full class sets, that is, a reader or textbook for each and every child. But even in the classes where each child had a book, the researchers observed very few lessons where these books were actually used. Children were certainly never allowed to take the books home. Teachers explained this with phrases like 'the standard of the material is too high' or books were 'too difficult for learners'.

So, if reading from books was not the primary learning activity, what was? The HSRC researchers found extensive use of worksheets — pages photocopied from various commercial workbooks that learners pasted into their exercise books or collected in files. The children did not get much practice reading extended texts and the sequence of worksheets often lacked coherence.

In the 160 classrooms, the researchers most often observed lessons involving teachers and learners reading aloud and reading in chorus. This occurred most frequently in Grade 1 but was not at all uncommon in the Grade 4 classes. The researchers speculated that the regularity of chorusing in the higher grades could be accounted for by the delayed shift to English as the language of instruction. As children did not understand the language, chorusing of English texts was a way to get around limited vocabulary. But even when Grade 3 children were engaged in reading tasks in a language that they understood, researchers observed that most reading involved children reading aloud together. Moreover, as children moved up the grades, the actual amount of time spent on reading, even whole-class reading aloud, declined. In only a few instances did the HSRC researchers witness learners reading silently.

Table 4.1: Reading opportunities during the observation period ordered from most to least common observation

1. Whole-class reading without the teacher *n* (valid) = 67	47 (70 %)
2. Teacher reading aloud to the whole class *n* (valid) = 65	45 (69 %)
3. Whole-class reading together with the teacher *n* (valid) = 66	39 (59 %)
4. Individual learners reading aloud to the rest of the class *n* (valid) = 68	38 (56 %)
5. Learners reading aloud together in groups *n* (valid) = 66	28 (42 %)
6. Individual learners reading aloud to the teacher who is monitoring them for guided reading *n* (valid) = 66	28 (42 %)
7. Teacher leading reading with class following silently *n* (valid) = 66	21 (32 %)
8. Individual learners reading aloud to others (partners/group) *n* (valid) = 66	20 (30 %)
9. Individual learners reading aloud in pairs *n* (valid) = 66	14 (21 %)
10. Individual learners reading silently in independent reading *n* (valid) = 67	13 (19 %)

Note: *n* stands for the number of observations
Source: Adapted from Reeves et al (2008)

When learners were reading aloud, what were they reading? The researchers noted that in many classes the most common reading activity was of an isolated set of simple words or short sentences: '[i]n more than half of the observations, *no* learners were required to read more than ten words. In more than half the observations *no* learners were required to read more than one or two sentences' (Reeves et al, 2008:134). The researchers almost never observed learners reading storybooks or non-fiction prose writing.

Macdonald (2002) characterised this instructional practice — the whole-class reading aloud, the choral reading of short words and sentences — as the 'rote-rhythm method'. Given the large classes and the poor training, Macdonald speculated that the underlying purpose of this practice was to reinforce hierarchy, with the teacher as expert and the child as novice. In the chanting ritual, children are forced to pay attention to what the teacher says and how she says it, rather than understanding what is being taught.[4]

4 In her later work, Macdonald has traced shifts in the core instructional practice in the late 1990s. From her fieldnotes on a Grade 1 classroom in the Free State, Macdonald wrote: The teacher, full of enthusiasm for her version of C2005, showed me a set of pristine, totally unused English language books (developed by experienced language writers), which she said that she did not use because they did not follow her Programme Organisers, and she proudly said she was making her 'own course'. The Programme Organiser for the fortnight was 'Transport', so the teacher had carefully labelled two pictures, *yacht* and *anchor*, on the board, and was drilling the children in reading them. Besides the fact that the phonic principles implicit in these two words are very advanced, it was highly unlikely that the children had any common frame of reference for, or experience of these objects (Macdonald, 2002:111-141).

This same pattern of reading, a handful of words or short sentence, was equally evident in numeracy and life skills subjects. In other words, by the time these Limpopo schoolchildren reached Grade 4, their experience of reading was mostly whole-class reading aloud of simple words.

A review of the literature reveals that the pattern discovered by the HSRC researchers in Limpopo is widespread. A few years earlier, Taylor and Moyane's (2005) base-line study for one of Joint Education Trust school improvement projects described almost identical practices. The JET study included 151 Grade 3 classroom observations in 24 schools. The JET team observed lesson after lesson, structured around written words or sentences on the chalkboard. Teachers would ask the whole class to read words written on the board, and then repeated this task two or three times, moving from the whole class to individual learners. The study described a typical Grade 3 class as follows:

> The teacher would introduce the passage with some Q&A and then read the pieces first, asking learners to read once, the process would be repeated, still as a whole class. The teacher would then call on one learner after another to read or one group at a time, following again by the rest of the class reading the same piece together. At the end of this part of the lesson, the teacher would write up a number of vocabulary words on the board and the class would chorus them together after her or him. The same words were then often given as class-work to the learners, e.g. 'write these words into your books' or 'complete these sentences filling in the missing words'. (Taylor & Moyane, 2005:38-41)

In about a third of all classrooms there was little evidence of any writing and another half of the observed classes did writing that consisted of one word or short sentences for about 15 minutes. There was very little evidence of extended writing, that is, writing a paragraph or more. Over the year, two-thirds of the learners had completed less than three pieces of extended writing. When children did write, it was copying from the board or writing one-word answers or short sentences in the worksheets.

Sarah Howie and her team from the Centre for Assessment and Evaluation at the University of Pretoria are best known for their research on learner achievement, but as part of their studies they have provided invaluable insight into this standard instructional practice (Howie, 2008). In the 2008 study, of the more than 60 Grade 4 teachers Howie's team interviewed, most described using reading aloud to the whole class every day or nearly every day as the primary method of teaching reading. In Howie's words 'this is the most prominent reading activity listed by teachers'. Half of teachers in Howie's study admitted to the interviewers that their learners never read silently, and a further 30 per cent said that they only read silently once or twice a month. (Internationally about 60 per cent of learners read silently every day or nearly every day.)

The qualitative case studies (Gains, 2010; Bizos, 2010; Pretorius & Mokhwesana, 2009; Maphumulo 2010; Maswanganye, 2010; Pretorius & Currin, 2010) confirm and extend our understanding of the basic pattern of reading teaching practice in South African disadvantaged schools. Pretorius and Mokhwesana's (2009) in-depth study of a single, poor (Quintile 1) primary school in a township on the eastern side of Pretoria explored the complexities and dynamics of changes in teaching reading in Northern Sotho. In the context of this base-line component of the study, they documented the teaching of reading prior to the onset of the project in the following way.

> Reading was primarily taught from the chalkboard, where lists of letters, words and sentences were written up and the children were drilled in identifying and saying them aloud in unison. Writing was done in worksheets pasted into exercise books. Particularly striking was the absence of storybook reading in the classes. The teachers did not read stories to the class, and there were no little texts or books for children to read individually or in small groups to practise reading in meaningful contexts. There is a N Sotho reader for the Grade 1s (comprising 392 words) but copies of the book were stored away in classroom cupboards because it was felt that the book was too difficult for the learners. Thus, during the first year of formal schooling the children had no opportunity to practise reading a connected text in N Sotho of less than 400 words. This is in sharp contrast to effective schools where good Grade 1 readers read up to 1,200 words per week. (Pretorius & Mokhwesana, 2009:60)

Bob Maswanganye's (2010) study of three primary classrooms in the Morelele region, a rural area about 60 kilometres north of Pretoria, reveals something of the residuals of earlier reading innovations. The three rural schools in Maswanganye's study varied in size but shared a basic rhythm of instruction. The school day always began at 07:45 with morning prayers, and singing a hymn and the National Anthem. Inside the classrooms there was a little variation. In one of the schools 'there are no bulletin boards and thus no pictures or sentences for accidental reading … There are no resources in the classrooms, except the chalkboard and the chalk'. Another had 'the schoolbooks supplied by the department … neatly packed in boxes in the computer laboratory. The latter is not used'. Overall, however, Maswanganye characterised all the schools as 'print-poor environments'. The teachers all shared a deficit view of learners, attributing weak reading skills to 'laziness'. Observing classrooms, Maswanganye found teachers predominantly using 'reading aloud' and 'choral reading' methods. One teacher described how he had his learners read aloud three times in unison (choral) before he got the more 'able' readers to read individually. A second teacher described a slightly different approach. Having been trained by Molteno, the teacher began her reading lessons by explaining the words to the children in their home language and then got the learners to point at the words that they were reading and

chorus them. The teacher believed that 'pointing' ensures that children remained focused on the words. Despite the importance the teacher attributed to the method, she acknowledged that although learners paid attention to the words they were expected to read, they were unable to read when requested to do so individually.

But what about the teaching of phonics in African languages? Thabisile Maphumulo (2010:1-82) was specifically interested in how Grade 1 teachers teach phonics in isiZulu. Before observing their practice, she asked teachers about their methods. This is how one teacher described what she does:

> I first teach or drill 'imisindo' (sounds) dominating in the story for example 'p'. We create words using the letter 'p' for example ipeni, upopo, etc. We read the words, create our own sentence using these words, we then read the story to demonstrate to the learners how to read and then allow them to read as a class and then as individuals.

The teacher's description of her approach suggests the lingering traces of the influence of various education NGOs, but without a systematic structured programme, it is likely that appropriately sequenced sounds and blends will be neglected or glossed over.

Nadia Bizos's (2010) research uncovers additional aspects of the prevalent instructional practices. While much of her study concentrates on what she called 'the worlds of six readers', gifted young people who become active readers in their personal lives, her study provides new insight into the meaning of in-school reading events. These events included reading aloud in assemblies, reading in class for information retrieval, and reading activities that took place during the Drop Everything and Read (DER) period.

The case study school, which Bizos has called Umholi, was not a particularly poor or under-resourced institution. It had a modest, if ageing school library stocked with books supplied as part of the READ/Business Trust *Learning for Living Project* in 2001. Despite having a variety of reading resources, including non-fiction books, old encyclopaedias and a few volumes of poetry, Bizos observed that the school made use of only a very limited set of texts for reading.

Within the school, Bizos observed that reading-aloud events, surprisingly, were a key component of the life of the school. The small group of good 'readers' were often called on during the school year to read, perform dramas and present poetry in class or at assemblies. Bizos points out that by identifying them for special reading duties, the school principal split the learners into two groups—an elite that could and would read, a group he called 'the readers', and the rest of the school population who would not be expected to read, certainly not to read at assembly. The weekly reading events at the assembly were clearly designed as a way to demonstrate the school's commitment to reading, but as Bizos points out, they constructed reading in a particular way; reading was a public performance. In Bizos's words:

> Divorced from any meaning-making, the children take up this 'performance' discourse and reading comes synonymous with 'reading aloud' and 'practicing'. Concern with understanding what is being read is framed solely in terms of display or being seen to perform well. (Bizos, 2010:188)

Teachers consciously teach reading as a public performance when they coach 'the readers' on punctuation and correct pronunciation. While Bizos focused on the reading aloud at assemblies, she noted that other South African research had also pointed to teacher collusion in the conscious or unconscious selection of only the strongest readers to carry reading in class. While reading as public performance was the primary way in which reading was enacted at Umholi Primary and was the focus of the teaching of reading, Bizos documents the reading pattern of the outlier learners, the 'readers' who actively read books and other material for information. Bizos's insights about reading as a public performance have been confirmed in the findings of other scholars researching the social practices of literacy in other institutional settings; such as the family (Stein & Slonimsky, 2006).

If Bizos's research shows that some children transcend the prevalent reading practices to become active and engaged readers, Paula Gains's study, on the other hand, reveals that reading teaching practices do not always mirror the historical patterns of the schooling. Following the standard pattern of qualitative case study research, Gains's study of reading teaching is a rich and dense description of the life in a number of Foundation Phase classrooms. Of particular interest is one of Gains's case studies, Ms Ndlovu's classroom in a former Model C school. After seven visits to Ms Ndlovu's classroom, Gains observed that the main teaching method that was used involved teacher-directed questions and answers with the whole class reading out aloud in a choral style (Gains, 2010). This is a brief example of dialogue in the classroom.

> MS NDLOVU: *Right, let's read together. All of us. You read after me. Read after me. Remember I told you that there are full stops and question marks and exclamation. You can't just read and read. You must stop somewhere. Take a breath, angithi?*
> LEARNERS: *Yes.*
> MS NDLOVU: *Yes. You read after me. You read after me. (reading) Mark, Mark.*
> LEARNERS: *Mark, Mark.*
> MS NDLOVU: *It is on!*
> LEARNERS: *It is on!*
> MS NDLOVU: *Come! Come!*
> LEARNERS: *Come! Come!*
> etc. (Gains, 2010: 179)

Gains suggests that while the reading-aloud, choral-reading approach she documents has serious limitations, it has the advantage of allowing weaker learners to benefit from listening to stronger ones. More generally, Gains argues, based on her interviews with teachers, that the prevalent instructional practice develops out of teachers' own schooling experiences, the kind of training they received as teachers, the resources at their disposal in schools and their perception of the children and what they are capable of doing. What Gains does not point to, but is evident in her data, is the extent to which systematic reading teaching practice in some instances has been displaced by the less effective choral-reading instructional practice.

Pretorius and Currin (2010) provide a good summary of prevalent reading practice:

> During the Foundation Phase (Grades 1-3) much of the emphasis in reading instruction in these schools is on the teaching of decoding skills (eg ba-be-bi-bo-bu; ma-me-mi-mo-mu etc.), the argument being made that these syllables form the building blocks of words in African languages (which are syllabic, agglutinating languages). There are not many storybooks or classroom readers in the African languages and schools are poorly resourced so storybook reading seldom occurs in the classroom. Reading is not typically associated with fun or pleasure, the reading of extended text is not a common classroom activity, and the development of writing in conjunction with reading is usually neglected.
>
> In many of these classrooms it is assumed that if learners can decode they can comprehend, so very little attention is given to reading comprehension … For many learners the transition from decoding syllables or words on a chalkboard, to meaningful reading activity involving extended texts does not happen easily. When the change to English as LOLT occurs in Grade 4, the learners are typically ill prepared to cope with reading in an additional language. Furthermore, as learners progress through the senior primary school phase, they increasingly encounter the more complex language of expository English discourse in their content subject textbooks in a language many barely understand. Reading as a tool for learning ('reading to learn') is thus never properly developed, in either the home language or the LOLT. (Pretorius & Currin, 2010:68)

What do all these research studies together tell us about the prevalent reading teaching practice in rural, township and a growing number of suburban schools in South Africa? While the dominant language practice continues to be initial teaching in an African language with a transition to English in Grade 4, there is a significant number of schools attended by non-first language English speakers where English is taught from the start. Whether teaching in an African language or using English exclusively, similar patterns of instruction are evident across a spectrum of schools. The features or characteristics of the practice include extensive use of the reading-

aloud and choral-reading method, that is, the rote-rhythm method, and some phonics teaching that focuses on letter names and sounds, with little emphasis on blends, encoding and decoding. It includes limited exposure to vocabulary in both the Home Language and First Additional Language that are often taught as single words or short sentences. Writing tasks predominantly involve copying from the blackboard, or filling in single words or short sentences in photocopied worksheets. While reading resources may be available, few classrooms have whole sets of appropriate graded readers either in English or African languages, and as such few children get regular experience reading extended texts at the appropriate level. Within classrooms, teachers attribute the slow progress that learners are making as readers to deficits in the children, although an elite group within disadvantaged learners are chosen to become the proxy 'readers' at events such as school assemblies or for outside visitors.

Some researchers may suggest that the instructional practice anticipates a particular social practice of literacy in the wider society (Stein & Slonimsky, 2006). That social practice of literacy has dual foci: rudimentary decoding for minimal levels of functioning in a modern society, and a symbolic meaning closely associated with reading as public performance or ritual. To make sense of the relationship between the prevalent practice and literacy in the wider society, it is necessary to locate the practice historically and link it to structures of power and privilege.

Structural conditions (ie teachers' initial training, large class sizes, limited availability of quality resources) and ideological frameworks (ie teachers' own conception of reading and beliefs about the capability of learners) together serve to maintain and perpetuate the prevalent practice. While it may be tempting to see the practice as a loosely coupled set of behaviours and resources, the research shows that the behaviours, beliefs and patterns of resourcing are better understood as a relatively stable, even entrenched, system.

Systematic reading instruction

My experience of the instructional practices that middle-class children receive comes from observations of the reading learning experience of my own children, in a former Model C school, who entered the Grade 1 class with considerable pre-reading skills and book knowledge. The teachers in the Foundation Phase at Parkview Junior clearly understood that the primary objective of the Foundation Phase was to ensure that all children learn to read fluently. Confident in their handed-down approaches, the teachers resisted the more absurd aspects of *Curriculum 2005*, but incorporated sufficient surface features to feel comfortable that they were implementing policy.

At the centre of the school's reading instruction was the Ginn Series (a 40-year-old Canadian collection of graded readers) that was located in neat shelves lining the classrooms. The Ginn Series was sorted by level, identified by different colours.

It was evident even to a casual observer that the classrooms had over 1 000 individual graded books. Some of the books were purchased in the 1970s, still bearing the stamp 'Property of the Transvaal Education Department'. Some were new titles within the overall Ginn Series; all, old and new, were covered in thick transparent plastic.

Along with the phonics programme and the 'look and say' words, the Ginn Series was the backbone of the instructional practice. The instructional practice began with differentiation — the teachers assessed children at the beginning of Grade 1, and subsequently assigned them to a reading group. The children then systematically worked through the various titles at their assessed level, and once they had mastered books, would move on to the next level. By the 1990s, the teachers described this as 'guided reading'. More than just organising instructional practice in the classroom, the Ginn books would come home from Monday to Thursday for homework. The teachers required a parent signature on a chart linked to the Ginn title. This would mean that the child would have read the book in group-guided reading activities in class and again with parents in shared reading at home.

While the Ginn Series was central in internal differentiation of reading levels within the classroom and formed the core of the reading practice at home, the instructional practice had many other components. The teacher taught phonics, initially letter-sound relationship, using commercial packages such as *Letterland*, but rapidly shifted towards more complex phonics such as blends, for example, (sh) (ch) (br) (gr). The phonics were also carefully graded from basic alphabetic principles and letter-sound relationships to more complex and less common blends from Grades 1 to 3. Along with the phonics teaching that took place in the classroom, Monday to Thursday homework would routinely consist of exercises of various kinds that would consolidate and drill the phonics lessons of the day.

Given the large number of irregular words in the English language, along with the graded Ginn books and the phonics teaching, my children would routinely bring home lists of words to be learned as whole words (Dolch, or 'look and say' words). These were simple but irregular words whose structure did not follow the normal rules of spelling. These words were to be memorised.

While the Ginn Series, the phonic teaching and the memorisation of the 'look and say' words together formed the basic components of the instructional practice, the classroom teacher and the school as a whole used a range of additional texts, books to be taken home and read that were not part of the Ginn Series. These included additional books for both shared reading and independent reading at home. By Grade 2, my youngest son was bringing home books like Roald Dahl's *Fantastic Mister Fox*. Having seen the cartoon film and having heard the story on CD did not detract from the power of the language, and the joy of having it read aloud to him over and over again. All Grades 2 and 3 children would be expected to go to the school library and take out books. The school's library would have a wide range of classic children's books. Favourites included *Where the Wild Things*

Are, The Hungry Caterpillar and multiple copies of Dr Seuss's 236-word *The Cat in The Hat* and *Green Eggs and Ham.* The school would also make use of specialised classes for children who appeared to be falling behind or not making sufficient progress.

Reading instruction would change dramatically in Grade 4. From a strong belief that the teaching of reading was the core business, reading was something children did in the process of learning in the subjects. Reading in English was reduced to short paragraphs embedded in worksheets, with only one or two books, such as *Charlotte's Web*, assigned to be read during the year. This pattern of teaching reading would continue until Grade 7, by which time three books were to be read as a class: *Harry Potter and the Sorcerer's Stone*, *Holes* and *The Silver Sword*. Children would not necessarily even read these books, as summaries and guiding questions were provided. Although a handful of children struggled to learn to read fluently, most managed to master the mechanics of literacy and gain a degree of automaticity in their reading.

By Grade 4 a small group within the class began to assume a new identity as active readers. To join the virtual club of readers, children would discover the pleasures of books and make reading a major part of the leisure-time activity. For many of the emerging 'readers', the *Harry Potter* books were the starting point. For many, it was not sufficient to have read the book and enjoyed it and taken on the activity as part of their lives, but it would begin the process of collecting books. Reading would become a regular pastime, sometimes a fixation occupying large number of non-school hours of the day. As in Nadia Bizos's (2010) study, only a relatively small number of children in these schools became 'the readers'.

Possibly the most insightful analysis of the instructional practice of reading teaching in middle-class schools is provided by Kerryn Dixon (2010) in her study of time and space in literacy teaching. For our purposes, a number of the moments in her study provide insight into the workings of this instructional practice. Dixon observed a teacher using the *Letterland* phonics programme in a demographically representative former Model C school. The teacher Dixon referred to as Helen had a clear idea that her children should not only learn the key letter-sound relations, but that 'correct' pronunciation was key. In a particular lesson, Helen was teaching the letter 'd'. Her emphasis was not only on the letter-sound correspondence but also on how the sound was to be produced. She asked a child in the class who appeared to be having difficulty identifying the sound. She repeatedly asked the class, 'What sound does Dippy Duck make?'

> HELEN*: What sound does Dippy Duck make? Does she say 'dee'? What sound does she make, Kurtis (boy's name)? 'd' — ippy 'd' — duck. What sound does she make?. . . When I say Dippy Duck what sound do you hear? I know that ducks make quack, quack but Dippy Duck is special like Annie Apple. Can apples really talk? [Class shakes their heads.]*

HELEN: *Listen, listen 'd'. Listen, listen 'd'. Say it. (Class says 'd' several times)*
HELEN: *When you say 'd' where is your tongue?*
CLAUDIA: *On your palate.*
HELEN: *Yes, it is on your palate at the top of your mouth. It's just behind your teeth. Put your tongue there. 'd' no, no, like that straight up you must feel it press up 'd', 'd', 'd', 'd' (class makes d sounds)*
HELEN: *Wow, so many dippy ducks. Thank you.*
(Dixon, 2010:80)

Dixon makes the point that the teacher appeared to be aware of the complexities of the linguistic utterance for second-language learners. Given that the structures of the sounds are very different in the dominant language of the children in the class (isiZulu), the teacher believed that it was imperative that the children not only learned the letter-sound correspondence, but that they learned the precise pronunciation.

While Dixon points to the socially constructed nature of pronunciation, there is an equally important insight about the nature of the instructional practice. The classroom event demonstrates the systematic nature of the teaching of phonics in the language of instruction and the additional instruction around pronunciation that ran parallel to the letter-sound relationship. But an equally important insight is that this teaching is taking place not in Grade 1, but in the Grade R class. The pace and coverage in the middle-class institution is substantially faster, and while Grade 1 teachers at the school would re-teach the entire phonics system again, with a strong foundation set for the majority of children the pace could be much quicker, allowing the teacher to move through key letters in the alphabet within the first few months of Grade 1.

The instructional tasks associated with the teaching of the 'd' letter-sound correspondence did not end with the addition of pronunciation; the teacher used the phonics lesson to develop vocabulary. The teacher asked the children to close their eyes so that they could all meet Dippy Duck together. Towards the end of the lesson, the teacher asked the class if they could recognise other words with 'd'. Once children had identified new words beginning with 'd', such as 'dragonfly', she initiated a conversation about those words.

HELEN: *Elwin have you seen a dragonfly before? Hey? Have you seen it, where did you see it?*
ELWIN: *By my cousin.*
HELEN: *By your cousin, where was it [joke] on your cousin's hair, where? (touches her hair)*
ELWIN: *By the river.*
HELEN: *By the river OK. Put your hands down guys. Let me tell you, dragonflies love to be around water. 'Cos you know why?*
[Vicky says something about eggs.]

> HELEN: *That's lovely Vicky but you put your hand up hey. They lay their eggs in the water. They live by the water because that's where they catch a lot of their food . . .*

The teacher then used another 'd' word generated by the children, namely, 'dance' and asked children to stand up and dance, suggesting the teacher's awareness of the value of multimodality.

The simple phonics instructional task associated with a systematic programme, *Letterland*, was used to provide opportunities to learn skills of pronunciation, extended language opportunities and was all done within the context of a tightly regulated classroom environment. The teachers in Dixon's school stuck closely to an older set of instructional practices that tended to emphasise skills such as phonics, pronunciation, vocabulary (language acquisition) and other related aspects in writing, such as spelling. But learners, particularly younger learners, had pleasurable experiences with books through the story time when they would sit on the carpet and listen to the teacher reading children's stories. Even in Grade 3 when there was an increased emphasis on formal school work, the teachers recognised the importance of a variety of experiences with varying genres of texts from conventional books to popular music lyrics.

In addition to the systematic introduction of phonics through the use of the *Letterland* programme in the pre-Grade 1 year, the school made use of what Dixon called 'indirect reading' instructional tasks. The teacher posted flashcards with a new word introduced every week. The teacher would use the simple words while the children were seated on the carpet, and the instructional activity was predominantly about memorisation. This is an extract from Dixon's fieldnotes:

> The children read the words stuck up on the wall as a class.
>
> Lisa (teacher) introduces a new word and asks if anyone knows what it is. This day it is *pot* and Tshepo puts up his hand and identifies it correctly.
>
> She sounds it out '*p-o-t*' and the children repeat after her. She makes some attempt to define a pot by saying: '*Mommy cooks with a* _____, *Daddy eats from a* _____', the children respond with 'pot'.
>
> She then goes around the class showing flashcards of the following words that each has to read: on, at, us, up, as if, in, it, is, am. (Dixon, 2010:103)

The teacher went around the class and got all the children individually to read the flashcard words. The emphasis in this instructional task was exclusively on correct decoding. Dixon recorded that not everyone in the class was successful and that the teacher needed to do considerable classroom management to complete the tasks. In the cases of children who clearly had not been able to do the basic decoding, she did supplementary instruction of various kinds. The teacher added complementary reading activities that reinforced the flashcard task — children identified flashcard

words in old magazines the following day. Dixon noted that while the exercise entrenched reading skills, it provided them with an opportunity to work with authentic texts, to get to understand how these texts worked and get pleasure from looking at the pictures.

Dixon described the Grade 1 teacher's practice of monitoring each child's reading progress. The Grade 1 teacher or classroom aide would listen to each child read from the first reading book. Children were either called individually to the teacher's desk or they were called in small groups. The children brought their homework books and readers with them. On being called to read for the teacher they would stand at her desk. She would take their homework book and open it to the back where a page was stuck in with a record of the title and number of books each child had read. This was signed by both the teacher and the parent each time the child read to them. The school referred to this as the Fun Reading Programme.

This regular component of the instructional core was not always pleasant for learners, but ensured that children read regularly every week and that adults (the teacher and the parents), were able to track and monitor progress in the number and titles read, and also how well the children's reading aloud was progressing.

The extensive use of phonics instructional tasks continued into Grade 3. Dixon described a set of tasks that began with the teacher introducing a new set of spelling words with the hard 'r' sound for the week. The words were put on flashcards and stuck on the board. The teacher pointed to each of the words with a stick and the class read the words in unison.

> part, starch, farm, tart, hard, bar, card, chart, start, charm, dart, hard, yard. (Dixon, 2010:110)

The teacher then asked the children which words they were unfamiliar with. Once these words had been identified, the teacher handed the flashcard with the unfamiliar word to groups of learners with the task of looking up the definition in the dictionary and making up a sentence with each of the words. The children had some difficulty with the dictionary task, in most instances not going beyond the first meaning provided in the reference book. Two days after this instructional task, the teacher followed up by using this worksheet with a phonics lesson on the 'bossy r', which covered the rule in the words on the flashcards.

Dixon's observation was that rather than moving towards more complex extended texts, the Grade 3 teachers became increasingly reliant on worksheets. Beyond the implicit learner surveillance assumed by discrete bits of texts that teachers controlled, the actual reading demands associated with the worksheet were very limited.

Romy Ganasi's (2010) study of six Grade 4 learners and their affective responses to reading provides important insights not only to the question that animated

10 October 2002

Bossy r-

Fill in the missing words

1. I received a birthday _____ as my present.
2. She was playing a _____ with her fingers.
3. The learners thought the sums were very _____ .
4. A _____ was shining at night.
5. The school _____ at 8h00.

star, starts, hard, card, harp

Make your own sentences with

tart dart farm yard

Draw a beautiful card

her study but the wider patterns of reading teaching in better-resourced schools. Although four of the six children she interviewed were second-language English speakers, most described participating in a range of reading activities in the school. These included being read to by the teacher (story reading), reading silently in the class (independent reading), group reading (shared reading) and listening to *Charlie and the Chocolate Factory* on the CD player (reading-while-listening). The children described the pleasure they got from visiting the school library and borrowing books. One of the key insights from Ganasi's study is the importance that children place on reading books that are 'easy', which highlighted the importance of having access to an appropriate reading level of texts so that children can feel a sense of mastery. Another important insight from Ganasi's interviews with Grade 4s is the way in which television shows and films tend to re-inforce and encourage the reading process. Children found particularly appealing those texts with material that they had encountered in other media. These included *Iron Man*, *Bratz*, *Hannah Montana* and *Ben Ten*. Familiarity with characters, story-lines and pronunciation all provided scaffolding to unfamiliar texts.

Exceptional classrooms and teachers, the practices that break with what would be expected, illuminate the standard practices. As with the choral reading that is often done in privileged schools, systematic literacy instruction can be found in a small number of classrooms in poor schools. Gains, in her (2010) Grade 1 literacy teaching study, presented a compelling case of Ms Jali who taught in a very poor neighbourhood in Soweto, in a run-down school, with a Grade 1 class of 46 children. Having received training and related materials from Molteno, specifically the Bridge to Literacy (BTL) resources, Ms Jali used an effective three-part structure for most of her lessons. Gains observed Ms Jali move her large class seamlessly from two

short whole-class tasks to an independent learning activity. In the whole-class tasks, Ms Jali began by reading a story or telling a story; the second whole-class instructional task centred on a phonics activity. For the independent task, children would complete a related exercise in their workbooks, or do independent reading, writing tasks or drawings that related to the earlier tasks and activities. The children alternated between working as a whole class, working in groups of around 15, or working on their own. At the core of Ms Jali's instructional practice was the systematic use of the workbook, the sentence-maker and systematic use of the 10 isiZulu BTL graded readers. While the purpose of Gains's study was not to measure reading achievement, she noted that by the end of Grade 1, all the children in Ms Jali's class could read the 200 words in the BTL Grade 1 programme, and had read between 4 and 34 stories.

Conclusion

There are a number of lenses through which to interpret the two distinct patterns of reading teaching practice in South African primary schools. One traces the elements of components of the literacy practice historically, understanding how the elements have been assembled over time. The assumption in this instance is that many of the township schools have a set of pedagogic elements that have changed little since the missionary schools, with an emphasis on mass instruction in choral reading and rudimentary phonics, and with little emphasis on reading fluency and comprehension in the early grades. In contrast, the middle-class instructional practice, with the benefit of substantially more resources, over decades has expanded from the original core around synthetic phonics instruction and the use of graded readers to include a range of other components, such as reading recovery and the use of 'authentic' texts. While the historical and pedagogic lens is useful it does not capture the full complexity.

Another approach is to identify a wider set of variables associated with the two distinct sets of practices. This has been a dominant feature of the school effectiveness research and, to a lesser extent, the school improvement literature. Researchers from these perspectives have consistently identified a list of factors or variables in the schools that they believe are associated with the bimodal pattern of learner achievement. These include pacing and coverage among others. The bimodal pattern of learner achievement has been confirmed by Spaull (2013), but earlier efforts to explain it tended to look for a set of discrete factors or variables (Fleisch, 2008; Hoadley, 2012). While these factors or variables are clearly important, they do not really get to the heart of the matter.

What we see is an interconnected set of behaviours, resources, beliefs and identities that are interwoven or interconnected. The distinct class-specific instructional practice is then best understood as a kind of social fabric that is tightly bound together in ways that are interlinked and mutually reinforcing. In the South African

context, we have two distinct literacy teaching practices: one closely associated with disadvantaged urban and rural working-class schools and the other associated with middle-class schools.

Understanding the practices as mutually reinforcing, interconnected and interwoven allows us to understand why the practices are so difficult to change. While one component might be removed or modified, the system has many other components or strands that remain interlocked, with the consequence that the basic practice remains. If the innovation is short-lived, the longer-term pattern returns to what it was before, with all the other mutually reinforcing components pushing it back into the original shape. And in many instances, the innovations are 'adapted' so that they fit with the existing fabric shape and form.

Chapter 5
Learning from research

Introduction

There are at least three things missing or underdeveloped in the literature on educational change. First, much of the writing in the field makes use of a simplistic theory of the social structural or 'social order' side of educational change. Second, this literature tends to underestimate the sharp or the hard end of education change beyond the bland notions of accountability, fidelity and political will, with little attention given to values such as tenacity, fortitude and sacrifice. Third, the literature seldom conceptualises education reform as a learning journey.

In this chapter I review research on large-scale literacy improvement initiatives undertaken in South Africa over the past two decades. In the general spirit of 'evidence-based' approaches, I pay close attention to the findings of 'what works' in South Africa. However, in the processes of reviewing the corpus of publicly accessible evidence, a few very startling trends emerged. First, despite the large amount of research and the very wide range of projects, there are few systematic or sustained conversations about what the research tells us about 'what works'. Rather, individuals and research groups have single-mindedly pursued their own agendas with little reference to the wider field or any attention paid to the lessons learned from other researchers within the South African space. Theoretical posturing acts as a curtain that divides scholars and programme initiators in the literacy field. Where robust evaluations have been conducted and published, they almost always revealed that the programmes or initiatives were 'effective' (with statistically significant results). But there is little evidence that the field was listening or integrating existing research findings. One cannot but get a sense that at least in the public domain, and likely even in the informal networks, little systematic or collective learning takes place.

In this chapter I track the research findings of education improvement initiatives that were designed to either improve reading achievement or change the way in which literacy is taught in primary schools. The list of interventions I have selected is certainly not exhaustive, and have not necessarily chosen them because of their relative success or failure, although I will suggest a few have had some impact on the instructional practices on a small scale. They have been chosen because they illustrate important themes about the particular nature of the instructional core on the one hand, and the theory of change on the other.

By scale and longevity, the series of interventions in primary literacy undertaken by two of South Africa's leading literacy non-government organisations, READ Education Trust and the Molteno Institute, were the most prominent. I will equally highlight small, often single-school interventions undertaken by university academics, including the work of Renee Nathanson, Alida Kruizinga, June Machie, Jean Place, Lilli Pretorius and Matseleng Mokhwesana, Sindisiwe Murahwa and George Hunt. These interventions focus on various aspects of literacy, from a narrow focus on encouraging teachers to model reading aloud (shared reading), to sophisticated and ambitious objectives of getting teachers to develop high-quality literacy materials in the children's languages and then build pedagogy around these teacher-developed materials. The interventions vary considerably in terms of the philosophies and theories that inform them, from the multimodal pedagogy in Pippa Stein's intervention to an indigenous language materials intervention. The interventions varied in size and duration, in terms of the focus on specific aspects of primary literacy teaching, the theories and philosophies that underpin them and they also differed in the types of evaluation.

Part of the explanation for the absence of a sustained conversation about 'what works' is that most of the studies have failed to grapple adequately with the prevailing instructional core, or to understand how and why it has come to be dominant system-wide.

READ

For almost three decades, READ Education Trust was one of the largest and best-funded education NGOs working in the field of literacy in South Africa. It was established in the 1970s as a vehicle to provide library books to Soweto schools, but evolved to providing comprehensive literacy support. Over time, the READ educational designers developed a literacy intervention model intended to improve both reading and writing in disadvantaged primary schools. During the early 1990s, the core READ model centred on the supply of quality reading books and other learning materials, in-service training for teachers and systems for classroom monitoring and feedback to teachers. By the late 1990s, the model had evolved to a 'whole-school' approach that included school and district leadership capacity-building.

In terms of the READ literacy materials, the programme supplied schools with a fairly extensive set of authentic children's reading books, including titles adapted from the New Zealand '*Sunshine*' *series*. The *Sunshine series* was not designed as graded readers, but authentic texts with strong stories and beautiful illustrations. Over time, READ began to commission and publish a whole range of children's books specifically for the South African market. The book series that was introduced into schools in the late 1990s, 'Sunshine in South Africa', was targeted at Grades 2 and 3 learners in six provinces. The intervention programme associated

with these materials included a 'starter pack' of 24 fiction and 12 non-fiction books, 24 alphabet books, 8 concept books and 20 big books. In most instances, the schools were provided with six copies of each title, the number designed to be used in shared and guided reading groups (Elley & Cutting, 2001).

Although providing schools with quality children's books was at the heart of the READ approach, the organisation recognised the need for training teachers to use these new resources. While the organisation originally subscribed to a strong 'whole language' approach, by the mid-1990s it had shifted towards a 'balanced approach' that integrated various forms of reading, such as silent or independent reading, shared and group guided reading. The READ programme incorporated aspects of phonics as they emerged within the context of teachers using the story books. While teachers received regular in-service training, the programme also made use of periodic classroom visits to teachers using the READ materials to check on programme fidelity and to provide on-site support and encouragement (Schollar, 2001).

Two formal external evaluations of smaller READ projects were published in 2001. The first was an evaluation of two reading programmes implemented in the Eastern Cape in the mid-1990s (Schollar, 2001) the other was an evaluation of, the project 'Sunshine in South Africa' which tried to assess the impact of the donation of 4 000 books from Wendy Pyle Ltd., from the New Zealand series by the same name (Elley & Cutting, 2001). The donated books were given to 22 schools in six provinces with the formal evaluation taking place in 1997. Both studies used quasi-experimental designs with selected schools receiving the treatment. Schools that got the interventions were not chosen at random but were selected on the basis of a pre-determined criterion, and the results of the intervention (or gift) were compared to a matching sample of 'control' schools.

Both the Schollar and the Elley studies found significantly higher gains in literacy achievement in the intervention schools compared to the matching control schools. According to Schollar (2001), the quantitative findings showed that in both projects learners made substantial gains in reading and writing 'equivalent to approximately two-year gains in a three-year time span'. In Elley's study (2001) the conclusion reached was that learners who were exposed to the 'book-based' programme on a daily basis showed rapid improvement. In particular, Elley found that the vocabulary and sentence comprehension of Grade 2 children using the Sunshine books increased faster than those of children in the matching group, with an effect size of .64 SD. Grade 3 learners in the project also improved, but not nearly at the magnitude of children in Grade 2. The children in the intervention schools did more and better written work, and could express themselves in more interesting ways.

The Learning for Living Project was one of the largest and most expensive of the READ literacy interventions initiated in the late 1990s, was implemented from 2000

to 2004. Supported by R153 million from the Business Trust, the project aimed to increase literacy, writing and other cognitive skills in the most disadvantaged primary schools, decrease repetition, train teachers to deliver the national curriculum and provide key resources to schools. READ Educational Trust was contracted to be the service provider for the teacher and manager training, which included providing learning support materials and ongoing classroom monitoring.

Of the method used in the project, the then head of READ, Cynthia Hugo, described it as a 'book flood' approach, which involved providing schools with 'high-interest, quality books in the target language and training the teachers in methodologies of a balanced language programme, to ensure that learners interact frequently and effectively with the books'. The focus of teacher training was on three components of the balanced language approach: shared reading and writing, guided reading and independent reading.

Once the selection of schools was finalised, 895 schools were included in the multi-year project. Unusual for projects in South Africa, the initiative had both an internal and an external evaluation. The external evaluation addressed three basic questions: was the project delivered as designed, were the project's goals achieved, and could the impact be ascribed to the project? The evaluation was based on intensive research in 50 of the 895 schools and an additional 30 control schools that were not in the project. Schollar and his evaluation team interviewed principals, teachers, did large number of classroom observations and tested over 2 600 schoolchildren.

On the question of whether the project was delivered as designed, save for problems with monitoring, he found that the schools received the learning resources and the training that was intended. The project schools were certainly supplied with far more training and resources than the control schools. On the question about the impact that the materials, training and monitoring had on learner achievement and retention, Schollar's (2005) evaluation showed that the intervention schools improved and were clearly doing better than control schools on a range of indicators, such as:

- observations of displays of teachers' work
- use of teacher-made materials in use in classrooms
- increased time spent reading by learners
- improved quality of group work
- increased use of comprehension-type questions
- popularity of the learning materials.

But Schollar found little evidence in the project schools of extended reading practices, especially sustained individual reading. And while the intervention schools were clearly doing better than the control schools, the quality and quantity of writing in both was very low. He noted that few learners could complete simple

sentence stems, and almost no learners had the capacity to do anything more than reproduce the simplest and most limited vocabulary or spell words with any degree of accuracy.

On the actual magnitude of gains, Schollar reported findings presented in Table 5.1. While some experts queried aspects of the study design, particularly the non-random sampling and the use of different pre- and post-test instruments, the message from this particular study was that the intervention school learners benefitted in both reading and writing.

Table 5.1: Cohort 1: Covariance 2000–2004 (Grade 3, 2000 to Grade 7, 2004)

	N	Writing	Reading	Mean Literacy
Project				
Baseline	1 049	3.2	17.5	10.35
Final	666	38.1	60.7	49.4
Change		+34.9	+43.2	
Control				
Baseline	678	3.8	18.5	11.15
Final	388	33.4	53.3	43.35
Change		+29.6	+34.8	+32.2
Covariant		+5.3	+8.4	+6.85

Source: Adapted from Schollar (2005)

The gains were evident in the group that would have been predicted to have made the greatest improvement. Children who had been part of the intervention for four years were clearly reading better than children in the control schools. But Schollar also observed that the gap between the intervention and control schools appeared to plateau in the four years with very little difference in improvement of writing and reading measured in the fourth year of the intervention. Schollar observed:

> Children still spend most of their time in classrooms, whether project or control, either speaking or listening, much less reading and very little writing. Although there are clear differences between project and control schools that are reported and discussed . . . , this situation remains generally true — education in the majority of South African schools remains essentially a verbally-based activity. (Schollar et al, 2010)

Deeply telling are Schollar's comments on the classroom observations. Over the life of the project, he observed that most teachers focused on oral competence and little attention was paid to reading and especially writing. Most of the classroom observations found the same basic activities: teachers reading stories to children who listened silently, following the teacher reading the text in the Big Books; teachers reading and children repeating in unison, sentence by sentence or paragraph by paragraph; and children reading aloud in unison in smaller groups. Schollar's field team almost never observed silent reading (possibly because teachers believed that it was better for the observers to observe them in an active role in the classroom). Schollar also pointed to the absence of writing, particularly extended writing in all schools. The one major shift in teacher-initiated learning activities was the substantial increase in the amount of time teachers in the project schools spent reading to children. This was made possible by the availability of fiction stories provided by the project, as these were resources not present in the control schools.

Embedded within the larger Learning for Living Project was a smaller study conducted by a group of US-based researchers (Sailor et al, 2010). While the core of the study continued with the theme of the effects of the 'book flood' model, it focused specifically on the impact that providing good-quality real texts in both the first and a second language had on children's reading in Grades 1 and 2. In the early grade study, the project provided each Grade 1 teacher with 20 Sunshine Starter home-language books, 20 big books, 120 'little books' (6 copies of 20 titles). The Grade 2 teachers got 40 Sunshine Starter English books, 40 big books, and 240 'little books' (again 6 copies of 40 titles); all the Grade 2 books were in English. As with the standard READ approach, teachers received training in the use of the materials and were regularly monitored and supported in the classroom. The project also had a range of add-on extras such as principal workshops and competitions.

The international evaluators were interested in three questions. First, did the intervention lead to high learner achievement?; second, was there higher performance in the select group of schools within the overall study schools?, and third, possibly the most important question given that the intervention focused on both Home Language and English, did improvement in Home Language literacy translate into higher achievement in literacy in English? The last question went to one of the most important questions in South African education—the rationale for teaching Home Language in the early grades has always been that learning to read in the home language strengthens a child's ability to learn to read in the second language.

The team of researchers developed Grade 1 literacy tests in all 11 official languages. The Grade 1 tests consisted of questions in the following areas: object identification (listen to a word spoken and identify the picture); letter/sound identification (listen to a letter sound dictated and identify it on paper); word identification (listen to a word spoken and identify it on paper); sentence comprehension (listen to a sentence and identify a correct picture); word production (learners read words

from a word list); sentence production (read aloud a sentence); and spelling (learners spell words). Save for the final task, most of the test tasks could only be administered individually. The testing was done in June with a random sample of 20 learners at each school site. The team analysed the data using Factor Analysis and Analysis of Variance to compare scores and subscale scores of learners in the control, select and intervention schools. Sailor, like Schollar, found that children in the 'book flood' intervention schools did statistically significantly better than the control schools in object, letter/sound, word and sentence identification in the first languages, but they did about the same on reading a sentence aloud (oral reading fluency). In English, learners in intervention schools did better in all subtests. Sailor concluded:

> There is a high level of value added to performance in both languages when a much more print-rich environment can be provided in the home language. These findings challenge the deficit myth often associated with children who come to school with a first language different from the medium of instruction. Further, the data suggests that providing instruction in both first and second languages can have a positive impact on development in both languages. (Sailor et al, 2010:36)

A number of questions emanate from the study. With no information on pre-test scores, how could we be certain that the gains can be attributed to the intervention and not some intervention bias? Even if the bias issue could be resolved, one of the data findings not discussed is the failure of the intervention schools to show statistically significant positive scores in oral reading in the Home Language. While children in the READ schools gained more than control-school learners in terms of their ability to identify words, sentences and letter sounds, they had not really moved ahead in fluency. On the question of language and language transition, Sailor did not provide new insights into the question about the appropriate stage for the introduction of English, even if the study did show that children can make substantial progress in English as they learn to read in their Home Language.

The variance of uptake in the project schools stimulated the University of Texas team (Sailor et al, 2007) to try to understand more about the schools that did make significant achievement gains. While the team's involvement originally consisted of evaluating learner achievement gains and measuring the fidelity of teacher activities, the secondary study made use of a school effectiveness approach. By exploring the sub-sample of the Learning for Living schools that had made big gains, Sailor and his team hoped to better understand what it would take to move the entire set of schools included in the project into a higher achievement bracket. Their study focused on seven schools. The schools had five features or characteristics in common. The schools put a high premium on creating a safe and secure environment, both in the conventional sense that children were physically safe and secure from the violence and intimidation that characterised the communities within which

the schools were located, and the psychological sense. The latter was established by ensuring an orderly and predictable environment, something largely created by the school principals. The second features or characteristics involved the work of the principals. These principals were not really instructional leaders; their role was far more about managing and procuring additional financial and other resources, managing and enhancing community relations and ensuring shared decision-making. Only in a few instances did the principals in these schools become actively involved in teaching and learning issues. The third theme revolved around teachers. From the interviews, the research team found that teachers in these schools were not simply good at teaching reading, but were described as committed, caring and collaborative. But what stood out for the research team about these schools had been something less tangible: a general sense of shared commitment on the part of all, parents, learners, teachers and managers, to the 'institution', and that they had a sense of pride and purpose in everything they did in the school.

Two key unexpected insights that emerged from the study are worth noting. First, while these were clearly the academically strongest of South Africa's disadvantaged primary schools, the researchers observed very little writing instruction.

> Our observations suggest that the 'conception' of literacy focuses on reading and not on writing. We observed practically no instruction in writing that offered the learner the opportunity to create texts. We saw spelling and 'copying' but seldom were there 'authentic' writing opportunities. (Sailor et al, 2010:385)

The second insight, made possible because the US-based researchers come from a learning context with extensive assessment, was the absence of any real external performance accountability in these schools.

Molteno

The origin of the Molteno Institute's literacy intervention project goes back to Rhodes University in the 1970s. A prominent fruit farming family from the small Eastern Cape village of Molteno donated funds to Rhodes University for a centre to improve the teaching of English. The Centre initially developed programmes to teach English as an additional language, but the leaders of the organisation soon realised that the problems in township and rural secondary schools in English actually stemmed from weak teaching of both English and African languages in the first years of formal schooling. Having commissioned research on early literacy teaching, Molteno found a prevalence of 'chorusing of meaningless syllables and copying of words in lists' as the dominant educational practice. Molteno staff also realised that schools did not really intend to teach children to read until the end of the lower primary phase. Molteno set out to develop a programme that would effectively teach reading in

Sub A or Grade 1. They decided that such a programme would need to begin in the mother tongue. Borrowing from a London School Council programme called Breakthrough to Literacy, a programme that had been used in the House of Delegates in Durban in the 1970s, Molteno adapted the programme (both materials and intensive teacher training) and translated it into isiXhosa and isiZulu.

By 1988, the Molteno Breakthrough to Literacy had been translated into Setswana, Northern Sotho, Sesotho, Xitsonga and Tshivenda and was used by 14 000 teachers reaching over 2 million learners. In the late 1980s, the NGO and the Breakthrough literacy programme expanded into Botswana, Namibia, Lesotho, Zambia and Uganda. By the late 2000s, it had been translated into over 40 African languages.

The basic package, what Molteno called the 'Breakthrough kit' had seven components. The Teacher's Sentence Maker and Holder was a display device teachers used to store and display vocabulary cards, the cornerstone of building sentences. The word cards were sorted into nouns, verbs, adjectives and adverbs, prefixes and suffixes. As Gains and Mfulathela (2005) described:

> [To] compose the sentence–*ubaba uyapheka* (father is cooking) the learners will need to find the two roots *baba* and *pheka* and combine these with the three prefixes, *u, u* and *ya* in the correct order to form the sentence. (Gains & Mfulathela, 2005:72)

The purpose of the sentence maker was to allow children to understand the morphology as they could see that what they 'say' could be represented in sentences with discrete words, and that these words were made up of syllables, and that syllables were themselves constituted by combining letters. Children were given miniature versions of the teacher's sentence maker to give them experience in making their own sentences and building their understanding of sentences, words and letter sounds. The Word Store was a holder of vocabulary words that formed the basis for work with the sentence maker. Children borrowed words or parts of words from the word store as they gained in confidence and vocabulary. The materials kit also included four conversational posters that included the core vocabulary and key sentences and a set of phonics posters that combined to form a phonics frieze. The phonics posters included 'new' sounds in the context of a picture. A set of 13 readers incorporated the core vocabulary, becoming progressively more complex as children progressed through them. The readers were designed for shared and independent reading. The 'kit' also included a learner's book (that incorporated a variety of tasks designed to develop children's cognitive and language skills), and two exercise books. Gains and Mfulathela (2005) argued that the Breakthrough was built on a 'balanced literacy instruction' approach with aspects of both phonics and whole language. But at its core it made use of a Language Experience Approach (LEA) in which children learned to read from the sentences that they themselves generated in group discussions. The initial stage of learning to

read focused on 'look and say' with the sentences and words that were generated by learners.

While Molteno's early grade intervention gained traction in Southern Africa, it was not accepted without critique. Wendy Flanagan (1992; 1997) was one of the more strident critics of Breakthrough. She criticised the programme as not giving teachers access to the theoretical underpinnings of the programme and for its heavy reliance on teachers learning the 'Molteno' procedures. She also questioned the evidence base of the programme, saying 'in spite of the enormous private sponsorship of this Project, running into millions of rands, there has been little substantial evidence to demonstrate its effectiveness'. The most serious criticism Flanagan raised related to the requirement that *Breakthrough* teachers 'follow the method as set out in the Manual'. In her view, this reproduced the authoritarianism of the previous political order. Drawing on Michael Fullan's early work, she suggested that 'top-down' approaches to change have little impact and teachers would inevitably become resistant to teacher-proof innovations. Flanagan also criticised Molteno for insufficient emphasis on an adequate supply of authentic books.

Another one of the major challenges faced by Molteno came from the organisation's linkages with the former homeland education departments. Flanagan (1992: 35-45) associated the core aspect of apartheid education, differentiation with a pedagogical strategy, with Molteno's approach.

> The state- (apartheid-) approved Molteno Project (a literacy project) selects the African junior primary classrooms in which to operate on the basis of 'more able' children, who, in turn, are sub-grouped into further ability groups after a first stage (presumably children not in a Project classroom must simply remain illiterate.) This is but one of the many practical examples of the implementation of this kind of education policy and the use of that discourse in the entrenchment of a notion of difference in terms of superiority and inferiority.

Molteno went into a period of decline in South Africa as a result of these types of criticisms (Rodseth, 2002). While not directly responding to the critique, Vic Rodseth acknowledged the problems that Molteno and Breakthrough to Literacy had in the late 1990s and early 2000s. In particular, Rodseth noted that while Breakthrough had originally been given a role in the new departments, it was rejected by provincial governments because it was perceived to be 'not OBE-compliant' and thus unusable in schools. The core concern that underlay the rejection of the Breakthrough approach was that the materials were too structured and insufficient attention had been paid to a thematic and integrated approach that had come to dominate the interpretation of OBE. Rodseth also defended the project against a range of other criticisms including the charge that it endorsed an early transition to English in the Bridge to English and the concerns that the programme implementation was too prescriptive.

CLE literacy intervention

One of the few studies that used a quasi-experimental impact evaluation design was undertaken on the Concentrated Language Encounter (CLE) literacy development project (Condy & Donald, 2005). The project began as a pilot in 1998 with funding from Rotary International and the Western Cape Education Department and was rolled out to a large number of schools at the beginning of 2001. By 2002, the project was working in 173 schools in three districts with multiple teachers being trained in each school. One of the leaders of the project, Janet Condy, described the CLE approach as 'based within the whole language tradition and making use of principles of scaffolding and active learning pedagogy'. The model originated in Thailand and one of the technical coordinators travelled to that country for training.

In practice, the CLE intervention model was heavily oriented towards training. Teachers in the project schools were trained by a team of department advisors in workshops held twice a year in their respective districts. These departmental teams were themselves trained by the technical coordinators (and the authors of the evaluation report). The training included an overview of theories of reading embodied in the CLE programme, followed by 'demonstrations of the specific reading processes involved at the relevant stages and phases and discussion of pedagogic processes such as mediation, scaffolding, and peer group learning' (Condy & Donald, 2005:48). While training appeared to be central, the intervention also provided unspecified follow-ups in schools and starter books, a teacher's manual and a training video. The project began running in 2001 in the Western Cape and was progressively rolled out to a large number of schools by 2005.

The project evaluation used a quasi-experimental design. Project schools were chosen at random from the large population of schools and then control schools were identified that had demographic features that were similar to the experimental schools. For the purposes of the impact evaluation, four schools were selected in the experimental group, and four in the control group. The same set of experimental and control schools were maintained for the duration of the project, that is, until the end of 2005. Baseline data was collected in schools in early 2001, and the initial analysis suggested that the experimental and control schools were statistically similar on all the usual characteristics. The evaluation had both a formative and a qualitative component. To measure literacy, the evaluation used a variety of subtests, including a cloze reading instrument designed for Grades 3 to 7, a dictation subtest, also for Grades 3 to 7, a vocabulary subtest for Grades 1 and 2, and a language and creativity test for all children in primary school grades. In the endline test designed for all learners, the evaluators tested 801 children in the experimental group and 742 children in the control group.

The evaluators described the 2002 results as 'encouraging'. The intervention had only been operating for less than two years in Grades 1 to 5, but the evaluators found statistically significant positive results in the experimental schools on the dictation, language and creativity subtests. The highest gains in the experimental

schools were in the language and creativity domains, which the evaluators believed was educationally meaningful as it assumed strong vocabulary development and suggested better integrated language and thinking development as a result of the implementation of the CLE intervention.

Reading is FUNdamental

While most of the interventions described in the chapter covered hundreds of schools, the small-scale study of intermediate school reading is possibly the most illuminating. Led by UNISA linguist, Lilli Pretorius, the aim of the project was to test the effectiveness of providing a basket of literacy-related resources and teacher training. Specifically, the researchers were interested in the relationship between reading improvement in the Home Language and English First Additional Language (they refer to these as L1 and L2).

The intervention was piloted in a highly impoverished school (Quintile 1) in a large township west of Pretoria. While the school had a typical language policy — Northern Sotho as the Language of Learning and Teaching from Grades 1 to 3, shifting to English as the LOLT from Grade 4 onward — the researchers noted that the majority of the children, who identified themselves as mother-tongue Northern Sotho speakers, did not speak the standard Northern Sotho (Sesotho sa Leboa) but rather non-standard urban koine' known as Pretoria Sotho (Sesotho sa Pretoria). Pretoria Sotho, a mixture of Northern Sotho, English, Afrikaans and words from other local languages, was the dominant language in and out of school. It is also an oral language. Sesotho sa Leboa, in contrast, is primarily experienced in written form by learners in the school.

Prior to the intervention, the teachers identified a major problem with reading performance, but were unable to name the reasons for underachievement. The researchers observed that few classrooms could be described as 'print-rich environments', and although school children did receive textbooks, they were almost never allowed to take them home. The school had a 'library', which was little more than a storeroom-cum-office filled with old textbooks. In terms of teaching reading in the Intermediate Phase classes, the researchers observed little formal teaching, reading homework was uncommon, and learners were given almost no opportunity for extended reading or writing.

The intervention itself consisted of three core components: providing print resources, building teacher capacity, and building parent capacity. On the first component, the intervention built a library collection with a functioning computer, printer, scanner and library software. It also provided an assistant to manage the school library. The project purchased and catalogued new English and Northern Sotho book collections. Within two years, the collection included over 5 000 titles (but with only 170 Northern Sotho books). On the capacity side, the intervention organised teacher workshops every two weeks. These training sessions focused on

reading and information literacy. Training session topics included comprehension, vocabulary development, reading assessment and integrating the library into classroom practice. While the training initially focused on Grade 7 teachers, the teachers and the team extended the training to include all teachers. The researchers also arranged for regular (every term) visits to a 'twin' school to observe 'good' reading practice in the equivalent grade. The intervention also included a parent or family literacy component.

A number of key insights emerged from this study. First, the problem of language and education was found to be more complex than policy would suggest. There was a mismatch between the language that is spoken and the standard language that the school uses as the First Language or L1. Second, strong gains in the First Additional Language or L2 can have direct benefits for the First Language rather than the presumed relationship, where children gained benefits from strength in the First that was transferred to the First Additional. The insight came from an analysis that showed that the greatest gains were in English, but those learners whose performance improved mostly in English were equally likely to make gains in Northern Sotho. Finally, and this was possibly the most significant of all insights, at the beginning of a project all learners tended to perform more or less at the same level. The researchers found that the successful intervention was likely to increase the intra-school or intra-class variation. Successful teachers were thus required to work with a wider range of reading ability within a single class as some children improved their reading (Pretorius & Currin, 2010).

While the focus of the project was at the Intermediate Phase, a second part of the project was the work on teaching reading in the Home Language within the Foundation Phase. The shift towards the early grades came as a result of the research team's realisation that many of the reading problems in the middle years emanated from the poor teaching in the early years of school. The researchers recognised that simply by virtue of teaching in the medium of an African language there was not a guarantee of academic achievement. Their observations confirmed that the teaching of reading in African languages at the Foundation Phase was often very poor. Much of it focused on the drilling of letters and sounds and little time was given to proper phonics teaching and the many other key aspects of early reading, such as phonemic awareness, oral language development and vocabulary, word recognition and lexical access, and comprehension and wider book knowledge.

Like the pilot study, the subsequent FUNdamental reading project intervention consisted of building resources, both a school library and classroom collections, and capacity-building for both teachers and parents. To evaluate the impact of follow-up intervention, the research team assessed 10 learners twice a year over four years. Each assessment was extensive, including instruments that measured letter and sound identification, phonemic awareness, story recall, book behaviour, vocabulary, word recognition and extended writing. The assessors, all working in Northern Sotho, assessed learners one-on-one.

Over the period from 2005 to 2008, the project tested a random sample of Grade 1 learners twice a year and reported pre- and post-average scores on letter knowledge, auditory discrimination, phonological awareness, word recognition, story recall, vocabulary and book behaviour. Pre-test scores rose dramatically in letter knowledge and word recognition, signalling improved learning in the Grade R class. In Grade 1 overall, learner pre-test scores improved dramatically in storytelling, book behaviour and word discrimination. The children showed consistent pre/post-gain in vocabulary.

The researchers observed that over the three years the successive waves of learners who were assessed appeared increasingly familiar with storybook reading routines, such as listening to stories and retelling the stories afterwards. Successive groups of children were more familiar with how to handle books and the very idea that a book tells a story, and just generally more familiar with literacy activities. What the researchers found and were clearly very disappointed in was the limited growth of children's Northern Sotho vocabulary, which in their view was at least two or even three years behind what they would have expected.

They attributed the improvement in the range of literacy related skills, both decoding and comprehension, to real and substantial changes in teaching practices. While the classroom observations were not as systematically undertaken as the learner assessment, the research team assembled a considerable set of evidence including observation notes, interviews with teachers, photographs and other field notes. The researchers observed very substantial changes in the visual environment in the classroom, with a significant increase in print materials on the walls. They found a range of new resources in use, including teacher-made charts, flashcards in Northern Sotho, and the regular use of the small collection of Northern Sotho storybooks. And while the basic teaching of letter/sound relations continued to rely on chanting from letters and words on the blackboard, the incorporation of a range of other techniques, including flashcards, meant that decoding was generally taught better. But it was on the side of comprehension that the researchers observed the greatest change in teacher practice. 'Storybook reading in Northern Sotho has become a regular — and much loved — activity in the Grade 1 classroom' (Pretorius & Currin, 2010:67). According to the researchers, the teachers really took on the new skills of story reading, and used it extensively to introduce lessons and themes, to develop vocabulary and get children to enjoy the Northern Sotho language.

According to Pretorius and Mokhwesana (2009), the new knowledge, broader understanding of decoding as well as comprehension and vocabulary techniques, including story reading, changed teachers' expectations:

> . . . at the start of the project, one of the Grade 1 teachers expressed hope that by the end of the year some of her learners would be able to read some words. When it was suggested that many of the Grade 1s should already be reading short simple texts much earlier in the year, she expressed polite disbelief. After

> spending a morning at a high functioning school where she observed Grade 1 learners read storybooks in the second term, she was motivated to work towards similar levels of literacy achievement. By raising her expectations of what learners could achieve, she worked to that goal. . . . If the teachers were pleased that half the children were actually reading short little texts by the end of the year, this suggests that prior to the intervention they expected less than this. This small change in itself is a necessary step to shift teachers' expectations and hence learner performance over time. (Pretorius & Mokhwesana, 2009:67)

Another side of teacher change was their relationship with parents. The researchers observed much greater emphasis on parent-teacher partnerships, letters sent home about children's reading problems and constant reminders to parents to get their children to read to them. The researchers also noted a new discourse, a new language being used by teachers about reading, a language that filtered into conversations between teachers, between teachers and learners and between teachers and parents. Teachers' growth in their ability to be good storybook readers contributed to their professional confidence and overall sense of efficacy.

On the wider change knowledge, the researchers observed that the friendly relationship between the teachers and the project staff helped create a non-threatening space within which teachers could test out their new knowledge and activities they had learned in the classroom. The researchers suggest that a strong component of the success of the project was the encouragement, affirmation and the regular feedback that the teachers received (Pretorius & Mokhwesana, 2009).

Family Literacy Project

Located in a remote rural district in KwaZulu-Natal, the Family Literacy Project (FLP) was designed to assess the efficacy of a family literacy model. The aim of the model was to improve the attainment of four assessment standards as specified in the National Curriculum Statement. In terms of this standard, Grade 1 learners were expected to:

1. use visual cues to make meaning (eg use pictures to interpret meaning)
2. role-play reading (eg hold book the right way)
3. make meaning using written texts (eg retell the main ideas in a story that is read to them)
4. recognise letters and words (eg read simple materials).

To assess the effectiveness of the training of teachers in the family literacy project approach and the impact of the regular follow-up visits, the researchers conducted pre- and post-literacy tests in both intervention and control schools. The children

in the project classrooms made substantially greater progress than the children in control classrooms. The project model appeared to improve book behaviour, story retelling, isiZulu word recognition and isiZulu story comprehension. With regard to the children in the control classrooms who received regular teaching, Pretorius and Machet (2004) observed little evidence of progress in the acquisition of pre-literacy and emergent literacy skills. For example, they found that relatively few children could recognise simple high-frequency written isiZulu words such as *vuka, sana, ikati* and *ikhaya* by the end of Grade 1.

Additive Bilingual Education Project

Possibly the most telling example of the problem of learning resource utilisation is the story of George Hunt's (2006) Additive Bilingual Education Project (ABLE) in the Eastern Cape. The ABLE project was established in 2000 as a collaborative venture between a community school and a British university. The main aim of the project was to promote additive bilingualism for the first few years of schooling and a well-designed transition to English. One of the major obstacles to implementing isiXhosa teaching was the absence of appropriate Foundation Phase books and other reading materials in the language. As a result, the project emphasis shifted towards the production of these materials. With the assistance and direction of the facilitators from Scotland, the teachers, parents and learners set about using *ntsomi* (traditional stories about animals and supernatural beings) as the basis for writing and producing 28 teacher-made storybooks and posters pitched at Foundation Phase isiXhosa-speaking children.

In a follow-up study, a year after the project was completed and the materials had been distributed to teachers in the school, the researchers returned to the school to find that only one teacher continued to use the innovative reading storybooks. The evaluators suggested that at least part of the explanation for the 'failure to thrive' lay in the teachers' reluctance to really engage with the unfamiliar instructional materials. It also had to do with teachers' ambivalence about teaching in the mother tongue, particularly in the higher grades as they prepared schoolchildren for the Grade 9 examinations. But the main lesson from this study is that the availability of contextually appropriate materials did not guarantee their effective and sustained use.

Systematic Method for Reading Success (SMRS)

One of the most interesting, but not well-known, early grade reading intervention initiatives (Piper, 2009; Ralaingita, 2011) was conducted as part of a wider research venture to develop an international reading assessment tool. This early reading improvement initiative, called the Systematic Method for Reading Success (SMRS), was developed by Sandra Hollingsworth and implemented by the Molteno

Institute. Its focus was on children's learning to read in the first year of school in their home language. The intervention was designed to be implemented through 55 'lessons'. Each of the 'lessons' were intensive and tightly scripted. The 'lessons' were designed to be used daily for between 30 and 45 minutes. The assumption behind the approach was that learners need to be systematically introduced to letter sounds, blending sounds into words, recognition of sight words, learning vocabulary and comprehension skills, through a teacher reading aloud and reading words in decodable and predictable stories.

> SMRS operates on the premise that reading stems from teachers' systematically introducing students to learning letter sounds, blending sounds into words, recognizing sight words, learning vocabulary and comprehension skills through reading aloud, and reading words in decodable and predictable stories. Once learners master reading in the mother tongue, they will be able to transfer these skills to other languages. Key instructional materials include culturally appropriate, progressively levelled storybooks, and teacher read-aloud stories with comprehension and vocabulary questions. (Ralaingita, 2011)

Children began with Book 1 and proceeded to Book 25. The schools received two weeks' training and the books. All schools were monitored, specifically on the extent to which they completed the 55 'lessons' and were given assessment instruments to use to assess learners reading progress.

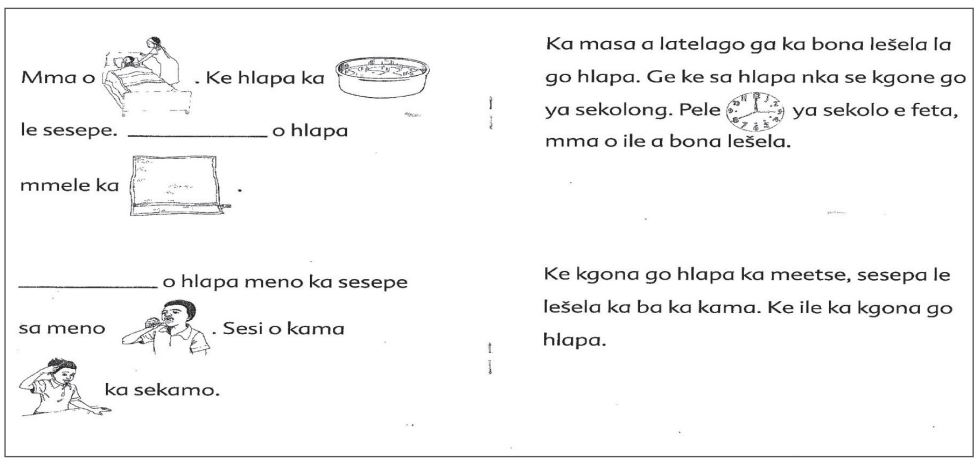


Figure 5.1: Systematic Method of Teaching Reading, Sepedi Book 16
Source: Department of Basic Education (2009)

The purpose of the study was to evaluate whether the method (SMRS) could be adopted by South African schools and if when used, the learners would perform better than their counterparts in schools using standard methods of teaching reading. The SMRS method was piloted in three languages—Sepedi, isiZulu and Setswana—in the North West, Limpopo and Mpumalanga in 2009. To do this, the study team selected a stratified sample of schools at district level that were receiving priority support from the provincial department based on their relative poverty and need. Ten schools were selected in each province to receive the 'treatment', that is, SMRS, and five were chosen as 'control' schools. Depending on the size of the school, between 10 and 20 learners in each school were selected at random for pre-intervention (January 2009) and post-intervention testing (June 2009) using the Early Grade Reading Assessment (EGRA) tool.

Teachers were required to teach each 'lesson' and complete the programme in roughly half a year. In practice, for a variety of reasons, the researchers found that on average teachers taught only 21 of the 'lessons' that they were required to teach. This meant that by the time the learners did the post-test, they had only completed half of the programme. While the focus was on using the structured programme, teachers and supervisors also were required to use a monitoring tool at every tenth lesson.

Despite the fact that on average teachers taught less than half the lessons required by the programme, the reading skills of the children in the treatment schools increased substantially compared to children in the control schools. The learners in the treatment schools made substantially (and statistically significant) higher gains in reading skills such as letter identification, familiar word reading, oral reading fluency and reading comprehension.

Table 5.2 shows that of the 137 learners in the control and 287 learners in the treatment schools, the percentage of children who could still not identify letters was dramatically smaller in the treatment group. In other words, the SMRS programme got most children with no letter identification skills and no ability to read familiar words, to master the skill in half a year with less than half the programme lessons taught. At the time of the pre-test, no children could read fluently aloud, but by the end of the programme, while half of the children in the control group had begun to read aloud, almost three quarters in the treatment group were reading aloud.

Table 5.2: Number (and percentage) of children who discontinued an EGRA subtask on the SMRS Study

Subtask	Measurement time point	Control	Treatment
Letter identification	Baseline	137	287
Per cent not able to identify letters		68.50 %	63.80 %
	Post-intervention	38	16
		23.30 %	4.20 %
Familiar word reading	Baseline	169	301
Per cent not able to read a word		84.50 %	66.90 %
	Post-intervention	68	88
		41.70 %	23.00 %
Oral reading fluency	Baseline	200	449
Per cent with no reading fluency		100 %	99.80 %
	Post-intervention	81	108
		49.70 %	28.20 %
Reading comprehension	Baseline	200	450
Per cent no comprehension		100 %	100 %
	Post-intervention	163	367
		100 %	95.80 %

Note: At baseline the number of learners in the control group was 200, with 450 in the treatment group. Post-intervention, 163 learners were assessed from the control group and 383 from the treatment group.
Source: Adapted from Ralaingita (2011)

Conclusion

Notwithstanding the Besharati (2016) meta-analysis of intervention programmes, what is so striking about the literature on intervention studies is the absence of a concerted and sustained conversation in which lessons from one intervention are learned by others. Protecting turf and the failure of NGOs to acknowledge weakness prevents the conversation. But equally, the smaller interventions are often driven by strong ideological orientations that do not lend themselves to genuine self-criticism.

In addition, there are a number of other general observations that can be made. Most published evaluation studies have positive outcomes. This could in part be

explained by the fact that academic journals tend to prefer positive findings (publication bias). There are a growing number of published evaluations that report statistically significant differences between test and control groups. However, few of them are genuine randomised control studies or have an alternative equally valid counterfactual design. Gains, when they are significant, tended to be on the lower-order reading skills. No study reports positive gains in extended writing in primary school grades. Where studies showed gains over the medium term (more than three years), the results plateaued rather than continued on a longer-term upward trajectory.

The published evaluation suggests that more often than not the basic literacy teaching practices remain relatively unchanged. These practices are characterised by whole-class, oral-teaching-directed activities. No studies have yet reported a substantial increase in the frequency of independent reading or extended writing in classroom observations. There is insufficient information about the pacing and sequencing associated with the staggered introduction of the two languages. While the study by Sailor et al (2010) shows that children can gain reading skills in Home Language and English if they have access to extensive storybook collections, many other questions related to language remain unanswered. Similarly, little is known about the relative efficacy of different components of the literacy teaching, for example, whole-class pedagogy compared to small group learning, authentic texts compared to decodable readers, and synthetic compared to analytic phonics approaches.

Chapter 6
Change at the instructional core

> To achieve large-scale reform you cannot depend on people's capacity to bring about substantial change in the short run, so you need to propel the process with high quality teaching and training materials (print, video, electronic). There is still the problem of superficial implementation when new materials are in use, and even new practices in evidence, without the deeper understanding required for substantial and sustained implementation. But you get farther, faster by producing quality materials and establishing a highly interactive infrastructure (crossover structures) of pressure and support. (Fullan, 2000:5-27)

Introduction

There can be no dispute about the very poor levels of achievement in literacy and mathematics in South Africa's primary schools. Research published in the past decade has only served to confirm the findings about the overall underachievement of South African schoolchildren and the bimodal distribution of that achievement. What we did not know half a decade ago and what we know now is that over the past 15 years reading achievement levels have largely remained unchanged despite a myriad of policy interventions. While elsewhere I have strongly foregrounded disabling factors such as poverty, ill health, expenditure and language as key factors in underachievement (Fleisch, 2008), this book has shifted focus to the proximal cause of underachievement — what happens inside schools and classrooms.

Having a deep understanding of the nature of the instructional core and its strengths and ultimately its limitations, is a precondition for change. An inadequate analysis of the problem can lead to ineffective strategies. For example, if we misinterpret the correlation between underachievement and low curriculum coverage, we might be inclined to develop interventions that stress close monitoring of curriculum completion. But more time and coverage, even focused time, is unlikely to change the kind of instruction provided and in turn substantially alter the patterns of underachievement. Simply more time chorusing or copying notes from the board is not what is required. These tasks, even if done more consistently and comprehensively, are only likely to have a modest effect, at best, on skills like oral reading fluency and comprehension. To improve academic achievement it is necessary to effect genuine change in teaching at the instructional core.

The first half of this chapter provides an account of the origins and evolution of the theory of system-wide educational reform and the turn towards theorising instruction. In the second half, I draw out how the theory has been selectively appropriated and adapted within a South African model of system-wide instructional change — the education triple cocktail.

I need to acknowledge that the theory and practice of system-wide educational reform elaborated here is not without critics. Pasi Sahlberg (2016), author of important work on Finnish education, refers to the approach as GERM (Global Education Reform Movement) and argues not only that it does not work, but that standardisation and narrowing of the curriculum has negative consequences for schoolchildren, teachers and the entire education system. Within the South African context, there have been similar criticisms of the local version. For example, Janks (2014) describes the use of standardised lesson plans as an approach that dehumanises the teacher, while Sayed et al (2013) see it as yet another uncritical policy borrowing.

What the research literature tells us about system-wide reform

Over two decades ago, two of America's most incisive educational thinkers, David Cohen and James Spillane (1993), developed a compelling framework for thinking about system change. They began by differentiating between educational policies in general and education policies that have the potential to influence and ultimately change instruction in classrooms. The latter they referred to as policies that offer instructional guidance, or more recently, the instructional infrastructure (Cohen, 2011).

They identified five policies that have genuine potential to change instruction:

1. instructional frameworks (what in South Africa we refer to as curriculum policy)
2. external assessment of learner performance
3. provision of instructional materials
4. monitoring of classroom instruction
5. policy requirements for teacher education and licensure.

While governments make education policy to regulate other aspects of education, Cohen and Spillane argued that it is only these categories of policy that have the potential to contribute to change in instruction. Their second important insight relates to how these five categories of instructional guidance, or what would probably better be described as instructional policies, are coordinated and implemented. The first category relates to *consistency*, that is, the degree to which the various instructional policies not only speak to each other, but are consistent and aligned. The second relates to the level of *specificity* or *prescriptiveness*. Instructional policies can be deliberately designed to be vague and general to allow for a wide variation of interpretation and adaptation, or they can be clear and detailed, specifying the what, when and how of teaching. Finally, there is the question of *authority and power*, that is, to what

extent government prioritises and enforces, placing its weight behind various degrees of implementation. This is sometimes referred to as 'political will' as it indicates the importance attached to government's commitment to use its implicit or explicit coercive capacity to drive change. The assumption Cohen and Spillane make is that when all five types of instructional policies are in place, when these are coherent and aligned, when they are tightly specified, and when there is a high degree of authority and power driving implementation, change at the instructional core is possible.

No one has done more to consolidate insights, experiences and case studies into a coherent body of change knowledge than Michael Fullan. In the work that he did in the mid-2000s, Fullan posited three basic propositions associated with the 'new' educational change knowledge. First and foremost, he argued that 'whole system reform' is possible if it is based on rigorous and robust change knowledge. Second, this education reform can and should be measured by the extent to which the education system's aggregate achievement levels are raised and achievement gaps between the achievers and the underachievers are narrowed. And third, change requires something akin to a 'perfect storm', the coming together of all the key change processes. At the core of Fullan's contribution is the recognition that systems need both inside-out and outside-in pressure and support, or in the more current formulation, accountability and capacity-building (Fullan, 2000)

Hargreaves and Shirley (2009) have situated the Fullan type 'change knowledge' in their Third Way system. Moving beyond the emphasis on progressivism, professional autonomy and individual innovation of the first wave of education reform, and school choice, teacher accountability, incentives and market-driven reform strategies of the second way, the Third Way of reform (a legacy of Tony Blair, Bill Clinton and Gerhardt Schroeder) was an intellectual middle ground that drew on new public management thinking and a strong commitment to public service delivery.

Hargreaves (2003) summarised the Third Way in 11 common features:

1. defining pedagogy and learning as the change focus for improved results
2. prioritising literacy and mathematics
3. setting system-wide ambitious targets
4. focusing on low-achieving learners to narrow the gap
5. prescribing tightly scripted programmes for teachers to follow
6. providing intensive teacher training on the use of the scripted programmes
7. introducing one-on-one coaching to support teachers and get them to persist with implementation of the scripted programmes
8. redefining principals' roles with reference to the new practice
9. getting teachers to engage with learner achievement data
10. aligning the improvements in teaching with the evaluation and assessment systems
11. involving parents and the community to support their children's learning within the specific context of the change approach.

Table 6.1: Features of the Third Way System Reform Framework

International tests used as catalysts	Change strategies often triggered by international test results or national attention generated on the weaknesses of the education system
Focus on literacy and numeracy	Basic knowledge and skills in reading, writing, mathematics and the natural sciences Narrowing of the curriculum, de-emphasis on arts and other subjects
National targets	Political identification of national numerical targets attached to the political cycle
External assessment and accountability	Sources of educational change are external innovations brought from the corporate world through legislation and national programs
Prescriptive curriculum	Tightly prescribed curriculum aligned to daily lessons with detailed scripting
Standardised learning materials	Provision of centrally procured and educative learning materials to all learners
Professional development	This takes a variety of forms from conventional short-course, just-in-time training to site-based coaching

While there was considerable debate about whole-system reform and going-to-scale in the early 1990s, the English National Literacy and Numeracy Strategy was a catalyst for the new 'change knowledge'. Prime Minister Tony Blair tasked Michael Barber to develop a comprehensive strategy to improve literacy and numeracy in all of England's primary schools (Barber, 2008). This he did. Within the first four years, the English numeracy and literacy scores appeared to have risen substantially. A few years later, a similar strategy was used in Ontario, Canada. But unlike the England example where scores flat-lined after an initial success, the Ontario literacy and numeracy scores showed strong and consistent improvement off an already strong base. While other research from a variety of contexts informed the finer details of the education change knowledge base, it was the experiences of England and Ontario that informed the basic assumptions of the new 'change knowledge'.

The first McKinsey & Company foray into the field, 'How the World's Best-Performing School Systems Come Out on Top' (2007), marked them as a player in the change knowledge space. The starting point for their report 'How the World's Most Improved Systems Keep Getting Better' (2010) was the recognition of how little was known about comparative system change and the basket of interventions that could lead to large-scale substantial and sustained improvement. During the

course of their consulting work and in interviews with leading policy-makers across the globe, the McKinsey team observed that few educational leaders have a 'mental map' about which kinds of interventions should be selected and how to put them together.

To fill this gap, the McKinsey team devised a relatively simple two-step study. Using Hanushek and Woessmann's normalised scale (based on a range of cross-national tests), they classified education 'systems' across the globe into four categories based on their performance as 'poor', 'fair', 'good' and 'great'. (At some points they also used the classification of 'excellent' for the Finland outlier.) Where consistent comparative longitudinal data was not available, they used national datasets to ensure the inclusion of a wider range of systems.

The authors identified from the list a group of 'sustained improvers'. These were systems that had experienced half a decade or more of improved student performance in international tests. They also identified a second group — 'promising starts' — in developing countries or societies in transition. Many of these systems did not have a long history of participation in the cross-national tests but had generated evidence that suggested they were improving.

Altogether, McKinsey identified the 20 most improved systems: 13 'sustained improvers', including some of the usual suspects, for example, Singapore, South Korea and Ontario; and 7 'promising starts', such as Armenia, Western Cape in South Africa, Minas Gerais in Brazil, and Madhya Pradesh in India. The McKinsey authors were careful to note that they were not studying countries, but systems, many of which were embedded in larger national contexts. This approach allowed them to compare small district systems with a small number of schools to national systems with hundreds of thousands of schools.

The second step in the McKinsey design was to collect and analyse the main interventions in each of these improving systems. While there were some common interventions in all the improving systems irrespective of whether they were moving from 'poor' to 'fair' or 'great' to 'excellent', the core finding revealed a differentiated basket of interventions associated with various stages in the change journey. Systems at the bottom of the hierarchy moving from 'poor' to 'fair', for example, Minas Gerais (Brazil), Madhya Pradesh (India) and the Western Cape (South Africa), used a cluster of interventions that included highly prescriptive mandated lessons (with associated student materials including workbooks), monitoring compliance through regular class visits, and setting performance targets based on universal external assessment. These systems also paid attention to the provision of basic classroom infrastructure. In contrast, systems at the top end striving to move from 'great' to 'excellent' adopted highly decentralised policies that emphasised teachers' and principals' individual drive for continuous improvement and innovation. These top systems stressed professional autonomy and placed a premium on peer learning.

While acknowledging the importance of context as a factor in the choice of interventions (such as the reluctance of Eastern European countries, given their recent history, to go with highly directive top-down interventions), the authors believe that the failure of most systems to make sustained improvement can be linked to a mismatch between their starting point and their policy intervention choices. Systems that currently have poor achievement levels have mistakenly experimented with policies that favoured high levels of teacher autonomy and unstructured peer learning. The McKinsey & Company authors add the caveat that correct intervention choices are a necessary but not sufficient condition for success. Intervention strategies need to be embedded into everyday classroom routines by strong district level support and consistent system leadership.

The McKinsey report needs to be located in Hargreaves and Shirley (2009) Third Way of educational reform. At a time in which quasi-market solutions appear again to be on the ascendance, particularly in the United States, the McKinsey & Company report clearly sees improvement as a function of the state or public sector (at whatever decentralised level). Building on insights from Michael Fullan, Michael Barber (the report's third author) and his England experience, and Ben Levin's work in Canada, the report extends what has been change knowledge of the geopolitical North into a global reform knowledge framework that now includes the 'South'. The 'South' incorporates a diverse group of countries from large emerging economies such as Brazil and India, to low-income countries, such as Malawi.

While acknowledging McKinsey & Company's contribution to the field, it is equally important to highlight some of its shortcomings (see Fleisch, 2011). This criticism notwithstanding, the McKinsey & Company cluster of interventions, again speaking specifically for countries that have reached near universal primary enrolment and a basic resource threshold on the journey from 'poor' to 'fair', the basket of interventions has a kind of *a priori* logic. Get the basics right, build sufficient school infrastructure and address schoolchildren's physical needs. Focus on those subjects on which all further learning is dependent, namely, reading and mathematics at the early primary phase. Put in place a simple and effective instructional practice and provide support for teachers as they become accustomed to it and it then becomes part of their daily routine. Monitor how well all the schools are doing using proficiency tests and help those schools that are not catching on.

What is not sufficiently developed are the theories of action and theories of pedagogy. At the core of the McKinsey change theory is the notion of collaborative practice, the importance of a mediating layer within the public service, and leadership continuity. The logic of the 'poor to fair' change journey would suggest that if a system pushes hard on new standardised scripted practices, a policy that encourages teacher collaboration is more likely than not to threaten lesson plan fidelity. In

systems at the 'poor' starting point in the change journey, the existing mediating layer (district and circuit offices) is more often than not a source of competing agendas. What is missing from the McKinsey theory of action is the role of civil society, and how non-state actors create pressure or demand accountability from working class and poor communities' social movements.

On the theory of pedagogy, the resource interventions, that is, provision of textbooks, structured lessons, proficiency testing and classroom coaching, need to be better theorised as part of an integrated system of instructional guidance or teacher learning. These resources, if they are well designed and carefully implemented, can drive teacher learning and real change in instructional routines. But equally, there is a real danger. Incorrectly calibrated lessons can further demoralise teachers. Scripted lessons, standardised testing and mechanical coaching can degenerate into authoritarian surveillance.

Luis Crouch (2011), as part of a larger analysis of institutional issues associated with the take-up of new instructional practices, suggests that one of South Africa's major barriers to change is midlevel functionaries and academics who simply cannot understand the need for simpler, more focused and more centrally directed approaches. He also points to capacity deficits and the absence of specialists with the knowledge of how to set up a clear and simple scope, sequence and pace of lesson plans, which teachers could implement. In his view, there is a list of key desirable actions that the government should take to address the literacy crisis. These include:

1. curriculum guidelines that specify the amount of time allocated to reading in the daily school timetable, reading resources that should be in every classrooms and standardised assessment practices
2. provide daily detailed lesson plans that focus on core skills to teachers who need assistance
3. provide storybooks for the most resource-poor schools
4. pilot external assessment tools in African languages
5. initiate randomised controlled experiments to test the effectiveness of mother tongue instruction and various methods of reading teaching
6. develop an advocacy campaign around early reading
7. provide reading kits for Grade R
8. government should prioritise literacy in the medium term expenditure framework.

In Crouch's view this cluster of interventions will do the job for a country that in his view is 'almost there'.

While Hargreaves (2003) acknowledges the value of the Third Way change

approach, his primary concern is to highlight the potential dangers associated with it. First, although it does address a limited problem, it does not translate into sustainable change. While scores may increase in the short term, systems reach a plateau and further gains are harder to achieve. Second, the focus on literacy and mathematics detracts from the other important school learning domains. Third, for some teachers, the prescriptive approach undermines their professional identity. But most important for Hargreaves, this approach can rapidly turn into a kind of 'cult' that requires blind obedience. Notwithstanding these cautionary insights, Hargreaves (2003:141) acknowledges:

> The scripted materials and strong support structures can benefit uncomfortable teachers at the beginning of their careers, uncertified or under-qualified teachers who work in poorer school districts, poorly paid and trained teachers in less developed countries, and other teachers whose knowledge, skills and overall expertise are weak or underdeveloped. A tightly driven programme of pedagogical change equips teachers with a repertoire of strategies that is inalienably theirs for life, and that can provide a strong platform for further improvement.

Within the education change community of scholars, there is a growing consensus about the general features of a change strategy for a system that is characterised by systemic underachievement.

The instructional turn

In the previous section, I sketched out the new thinking about system-wide reform and the consensus about the need for multiple interconnected interventions to drive change at scale. The other key idea, albeit more controversial, is the value of tight specification or prescriptiveness, particularly in the context of weak education systems. The third idea that needs clarification is change at the instructional core.

As outlined in Chapter 3, the starting point for an examination of the triadic concept of instruction and instructional capacity is Ball and Cohen's (1999:2) seminal paper.

> We focus on the interactions among teachers and students around educational material, rather than seeing curriculum alone or teachers alone as the main source of instruction. On this view, each of the three elements is essential, but instruction requires all three. Instructional capacity to produce worthwhile and substantial learning is a function of the interaction among these elements, not the sole province of any single one, such as teachers' knowledge and skill, or curriculum.

Hiebert and colleagues (2012; 2017) approach the conceptualisation of instruction from a slightly different angle. They talk about the importance of *teaching* rather than *teachers* as the centre of improvement. They point out that much of the American education literature tended to blame underachievement on teachers as individuals, implying that failure to learn is caused by individual teachers' lack of subject content and pedagogic knowledge. The conceptual shift from seeing 'teachers' as the problem to 'teaching' as the problem enables us to move away from what is simply in or not in the heads of teachers, to rather consider how poor learning outcomes are a problem of the practice stemming from the interaction of teachers with learners around educational materials. The key insight then is that improving teachers' knowledge alone is unlikely to be impactful. Upgrading the educational materials available in the classroom alone is unlikely to be impactful. These two essential elements of instruction are interdependent. Only if both are improved would it be possible to see change for real.

Although education-change scholars do not all agree on the definitions of the unit of instruction that should be the focus of improvement, that is, the academic tasks, activity or the lesson as a whole, they agree that improvement must centre on the smallest unit of instruction. City et al (2009) provide a strong rationale for focus on the smallest unit of instruction, the instructional task. They write:

> What predicts performance is *what students are actually doing*. The single biggest observation discipline we have to teach people in our networks is to look on top of the students' desks rather than at the teacher in front of the room.
>
> Accountability drives the task system in the classroom. As a result, students are especially sensitive to cues that signal accountability or define how tasks are to be accomplished. In addition, students tend to take seriously only that work for which they are held accountable.
>
> But to do what they are expected to do, they must know not only *what* they are expected to do but also *how* they are expected to do it, and what *knowledge and skill* they need to learn how. When we put teachers and students in situations where the task is vague and unspecified, but the expectations for performance are specific and high, we are expecting them to do the right thing without knowing the right thing to do. (City et al, 2009:30–31)

If learners' performance on assessment tasks, such as fluent oral reading and comprehension, are the outcomes we are most interested in and want to improve, we have to pay attention to whether or not the learners are routinely engaged in instructional tasks that provide them with opportunities to master skills and knowledge assumed in the assessment tasks. In Elmore's language, not only do learners need to know what they are expected to be able to do, but they need to be

given tasks that allow them to practise and advance their skills and knowledge of how to do that which we price in assessment.

For Elmore, accountability in education should be focused on the quality of instructional tasks or lessons rather than learner outcomes. Most politicians worry about test scores (or matric results), blame teachers or teacher unions, but pay insufficient attention to and hold the system accountable for the deficits in the instructional core, the source of the poor performance. Elmore again: '[i]n most instances, principals, lead teachers, and system-level administrators are trying to improve the performance of their schools without knowing what the actual practice would have to look like to get the results they want at the classroom and school level'. (2011:32)

So, if the instructional tasks or the lessons are the things that we should be concerned with and the prime driver to improved learning, how do we do this? The first obvious insight is that you need to change *teaching*, rather than *teachers* and that in order to change teaching, yes, you need to improve teachers' skills and knowledge, but also the educational materials that the learners use and the quality of the interaction around tasks learners do in the classroom. Elmore's final argument around the notion of the instructional core involves an idea that came to him from observing professional learning in the medical profession, particularly the learning that takes place in the hospital rounds involving doctors, medical students and patients: that improved teaching requires direct, face-to-face interactions in the classroom during the lesson and after the lesson. It is only by observing skilled teachers teaching the lessons with the educational materials, interaction that involves developing new language to describe and analyse the practice, being observed and giving feedback by experts, that improvement in instruction happens. Using his phrase, 'We learn to do the work by doing the work, not by telling other people to do the work'. We don't learn by listening to someone telling us how the work is to be done.

The education triple cocktail

Drawing on ideas from the National Literacy and Numeracy Strategy, instructional infrastructure and how teachers learn new instructional practices, the Gauteng Primary Language and Mathematics Strategy (GPLMS), initiated in 2010, was explicitly grounded on the emerging new change knowledge. The Strategy made use of multiple, overlapping and mutually reinforcing components of instructional infrastructure, all of which were tightly aligned, both in terms of their emphasis on instructional practice and in terms of the sequencing and pacing of the initiative's roll-out. Although the use of the Annual National Assessments was originally seen as an important pillar of the Strategy, in a low-trust atmosphere with national tests that were not comparable, performance accountability as a key driver of change

never really gained traction. The key levers of the intervention model turned out to be the prescriptive lesson plans, the provision of quality learning and learner materials, and face-to-face instructional coaching. Borrowing the name from the combination therapy that saved millions of lives and changed the course of the AIDS epidemic, the South African change model came to be known as the 'education triple cocktail'.

Prescriptive lesson plans

What are the key elements of the instructional infrastructure needed to overcome the inertia caused by the structural and socio-cultural, determined instructional practices? If the first stage in the change process is to re-engineer instructional practices (with the second stage to institutionalise the new practice), the catalyst for change was to be prescriptive lesson plans. The specifying of the lesson within the set of lesson plans was seen as the best way to get to the smallest unit of instruction. Hiebert and Stigler (2017:169-176) described it as follows:

> A place to begin understanding and improving systems of teaching, as with all systems, is to identify units for analysis that contain all the critical components of the system. Such units are ecologically valid, not too big to be studied, and repeat frequently enough to enable application of improvement methodologies. For classroom teaching, the smallest unit that preserves the key interactions within the system is the classroom lesson, a mini-system that repeats day after day, where teachers interact with students about content.

While there have been strident critiques of scripted lesson plans and principled objections to them within the teacher education community, this component of the instructional infrastructure is receiving growing recognition. Beatty (2011) reminded the research community that many of the key progressive education reform movements of the past, particularly Montessori's early childhood instructional programme, relied heavy on tightly scripted lesson plans to fracture old patterns of instruction and put in place newer approaches and methods. Reeves (2011), while only partially supportive of scripted lesson plans, showed how teachers moving into a new domain of teaching benefitted from designed scripted lesson plans to learn the new teaching practice. Hiebert and Morris (2012) have persuasively argued that detailed lesson plans, what they refer to as 'annotated' lesson plans, are a key resource or set of artefacts to guide the new practice that moves the debate from a focus on teachers to a focus on teaching. Shalem and colleagues (2016) have recently found that effective lesson plans might need to be more rather than less detailed to ensure that some of the complexities of the material are adequately communicated and that to really benefit teachers, their subject material knowledge needs to be enhanced substantially.

Detailed lesson plans are the blue-print for the new instructional practice. Not only do they assist teachers with segmenting or chunking the learning outcomes into manageable pieces around the 'lesson', they systematically sequence the learning tasks and activities in appropriately developmental ways. Well-designed lesson plans assist in maintaining learning momentum, pacing the work both within the lesson itself and across the academic year. Lesson plans that contain 'built-in' assessment tasks and instruments also ensure appropriate and relevant articulation between what is taught and what is assessed. One of the major strengths of lesson plans is that they can introduce new teaching techniques and methods in balanced and appropriate ways. Moreover, Hiebert and Morris (2012) observed '[a]nnotated lesson plans can store the knowledge acquired for improved methods of teaching, and common assessments can ensure that changes to lessons are improvements, not just changes'. For example, for Foundation Phase teachers who have never used systematic phonics programmes for learners learning English as a First Additional Language, prescriptive lesson plans provide a clear and simple guide — when, what and how — to introduce various bi- and trigraphs. These would be reinforced with linked vocabulary development oriented wall charts, 'look and say' word lists and graded readers that contain the words. To manage the novelty load associated with the initial stages of instructional change, the structured lesson plans provide teachers with new work rhythms and routines. Once in the new routines, which are a central feature of the lesson plans, teachers can more easily adjust to the new instructional practices. Finally, although prescriptive lesson plans greatly reduce the time required for mapping the daily lessons against the national curriculum requirements, the prescriptive lesson plans do demand that teachers spend considerable time preparing for the lessons, ensuring that they are familiar with the content of the upcoming lessons and that they have all the resources and materials ready at hand.

High-quality learner materials

The cost-effectiveness of providing quality educational materials has been documented internationally over the past 40 years and has been confirmed in a number of recent Southern African studies (Fuller, 1987). From the perspective of 'instruction', Ball and Cohen (1996) made a strong case for the centrality of learner materials and resources in large-scale change. They suggest that many education reformers have tended to disparage the use of textbooks or reading schemes because of an idealised version of teacher autonomy. As was common with the implementation of South Africa's experiment with outcomes-based education, Curriculum 2005, it was assumed that teachers would not follow textbooks but would make their own learner materials. This was a deeply problematic assumption as it often meant that learners went without the key tools for their academic development.

There is clearly a related problem in South Africa — that some of the key learner

resources/materials have never been available or even developed. For example, while almost all schools in middle-class communities have reading schemes or graded readers that allow learners to have extensive opportunity to practise reading with materials pitched at the right level, such resources have seldom been produced in African languages and therefore are not available to schools that have adopted African Home Language policies.

This learner materials problem has been recognised and, in part, has begun to be addressed with the provision of the Department of Basic Education's DBE Workbooks. While these resources are an important starting point, they certainly cannot and do not fulfil the full range of learning needs. It is clear that teachers require not only workbooks, but quality and appropriately levelled/graded readers, books for independent reading, decodable texts, high-frequency word lists, oral development oriented wall charts, and various kinds of manipulatives.

There are three key issues related to learner materials that need to be noted. First, schools often do not have a sufficient quantity of the high-quality materials to make them an effective component of the instructional processes. Although sharing between two learners may not be disadvantageous, it is often not optimal. Second, schools tend to view learner materials as consumables rather than permanent collections. This has been reinforced by the provision of the DBE Workbooks. This prevents a school from developing and implementing a medium-term procurement strategy, placing effort and care on using, maintaining and enhancing an appropriate and balanced learner materials collection. Third, Ball and Cohen (1996) have observed that learner materials are often seen as part of improving teaching and learning but are seldom understood as one component of a systematic approach to teacher development. There is clearly a need for tight alignment between learner materials and teacher development strategies.

Instructional coaching

Detailed, prescriptive lesson plans and high-quality educational materials are a necessary but not sufficient condition for large-scale change at the instructional core. Without teacher learning and teacher agency, these elements of the instructional infrastructure have little chance of transforming the everyday learning activities and tasks in the classroom. There is a growing recognition in South Africa, reflecting international trends (see Piper and colleagues' work in Kenya 2015a&b), that conventional forms of in-service training, which are short-courses, workshops and accredited academic programmes, fall short of transforming instructional practice. As such, new approaches to teacher development have emerged, many of which focus on the development of localised communities of professional practice. While teacher-led communities of practice may be appropriate for high-functioning systems where the primary concern is the sharing of best practice between teachers, this approach is unlikely to be effective in a low-function system that has few examples

of good practice that could be easily shared. And so, another main approach to teacher development, that is, instructional coaching, was included as the third component of the education triple cocktail model.

Instructional coaching is informed by a professional learning theory. D'Amico and Stein (2002) suggest that powerful professional learning needs to be *long-term*, *experientially based* and heavily dependent on *individualised* interaction between the novice and an expert in situ. To this, Kraft, Blazar and Hogan (2017) add the need for such development to be *intensive*, with interaction taking place anywhere between once a week and once a month, and *focused*, that is, working with teachers around the deliberate practising of specific skills. Through the one-on-one coaching process, teachers can acquire complex skills and knowledge and technical vocabulary associated with the new instructional practice. Instructional coaching is a social learning process that begins with (1) the novice observing a coach modelling the new practice, (2) the coach facilitating meetings and guiding the teachers in the use of the new practice, (3) observing teachers trying out the new practice and providing constructive feedback, and (4) setting up independence and ongoing peer-teacher learning spaces and opportunities.

While the prime purpose of coaching is to enable practice-based learning, there is a second and equally important role that coaching plays in instructional change. Since the seminal studies by Fullan (2007) and Hargreaves (2005) on the meaning of educational change, there has been a growing recognition of the significant impact, both positive and negative, that emotions have on educational change. There is an emerging body of research that suggests that instructional coaches can play a positive role in mediating the anger, insecurity and feelings of inadequacy that teachers experience as they shift their instructional practice. (For an example of this in South Africa, see Masterson, 2013.)

For instructional coaching to be effective, the coaches need to be carefully selected, thoroughly trained, and their role needs to be very clearly defined. There is evidence that in the absence of clearly defined roles, coaches increasingly take on administrative functions like distributing materials and collecting information (Fullan & Knight, 2011). Coaches who do not receive adequate initial training, sufficient ongoing support as well as careful monitoring and evaluation are unlikely to make an impact on instruction (Gallucci et al, 2010).

In the Gauteng version of the education triple cocktail model, the provincial department contracted outside NGO service providers to hire and manage the instructional coaches. McKinsey & Company provided pro bono support for the design of the coach training. At the height of the intervention (at the end of 2013), over 500 coaches were working in the project. Most of these coaches had experience as primary school teachers, and some had worked in teacher education. They were paid the equivalent of a school Head of Department. For many of the schools in the province, teachers in Grades 4–7 were subject specialists, focusing on either English

or mathematics. This meant that the second wave of coaches tended to be hired based on their subject specialisation rather than on phase (for example, early grade) expertise. Coaches visited every teacher in the programme at least once a month and in many cases twice a month. During their visits they modelled the new practice, observed and commented on lessons, and examined students' exercise books and assessment activities. In addition, the coaches also provided 'just-in-time' training to clusters of teachers, working through the scripted lesson plans a few days before they were due to be taught. They also helped to establish small 'communities of practice' for teachers who met every month after school.

But while the emphasis of the coaching process is on learning from and alongside an expert in situ (a process that has already demonstrated its usefulness in building trust), coaching also enables informal and unanticipated forms of internal accountability. In a system where teachers had actively resisted classroom visits by district officials and school Heads of Department, the GPLMS coaches made over 120 000 successful visits in the first three years, experiencing almost no opposition from teacher unions. This suggests that the coaching process was able to gain the trust of teachers. In the process, the coaching visits *de facto* opened up the classroom to external scrutiny. Opening up classrooms to outsiders and, by extension, the opening up of the actual new instructional practice to external appraisal, has enhanced professional accountability across the system.

Conclusion

By learning from international experience and global academic research, South Africa has developed its own unique instructional change model, the education triple cocktail. The model was constructed for and tested first in Gauteng Primary Language and Mathematics Strategy (GPLMS), a large-scale intervention that began in 792 underperforming primary schools in the Gauteng province in 2011. The following year the intervention was expanded to lessons across a slightly enlarged group of 'priority' primary schools. By 2014, at the height of the intervention's reach, 1 040 schools in the provinces were part of the intervention that included all Grades 1–7 teachers in three subjects — Home Language, English as First Additional Language, and Mathematics.

This chapter has explored the theory and practice that contributed to its development, particularly the assumption that system-wide instructional change, led by government, is possible; that such change would require closely aligned and interlocking components or instructional infrastructure; that the change needs high degrees of specificity or prescriptiveness; and that the focus must be on improving the quality of lessons and instructional tasks within lessons. The education triple cocktail and the theory that it draws from assumes that learning the new practice would require teachers to cultivate new professional knowledge and skills in

the right kind of emotional atmosphere. It assumes that this new professional capacity could only be mobilised if classrooms had more appropriate and better educational materials.

What the education triple cocktail model of system-wide instructional reform did *not* take up from the global education reform movement was its strong emphasis on standardised testing and performance accountability. Nor did it take up policies of education privatisation or school choice.

Chapter 7
The Reading Catch-Up Programme: a case study of change at the instructional core

Introduction

The international models of change knowledge and the broad descriptions of intervention programmes provide little real insight into the internal processes associated with system-wide instructional change. What are the actual mechanisms and dynamics involved in shifting the practices of thousands of individual teachers in classrooms in hundreds of schools spread over a large geographic area? What are the unique challenges of trying to make change at scale? To answer these questions, we need insights from information-rich case studies, for it is only in this kind of research that knowledge of the mechanism of change is likely to emerge.

During the middle of 2011, the first year of implementation of the Gauteng Primary Language and Mathematics Strategy (GPLMS), the management team began to recognise that it would be problematic to simply roll out the GPLMS programme into the Intermediate Phase, that is, Grades 3 to 7, in 2012. The evidence from the cross-national studies (Fleisch, 2008) showed that children who had not had access to the good-quality early reading instruction in the Foundation Phase years would be two or three years behind where they should be academically. Simply putting in place a new curriculum programme, designed for learners who did not have adequate foundational learning, would undermine the effectiveness of such a programme. The GPLMS management team therefore made a decision to develop and implement an intervention to remediate the learning gaps. This component of the GPLMS came to be known as the Intersen English Catch-Up Programme.

This chapter describes and analyses the design, implementation and impact of this instructional programme as a case study of the education triple cocktail. While it was unique in some respects as a systemic remedial programme, it contained the three key elements embedded in the GPLMS change approach, that is, prescriptive lesson plans, provision of high-quality learning materials, and just-in-time training and on-site coaching. From a research perspective, the Catch-Up Programme had two advantages. First, the designers wrote extensive documentation of all components and stages of implementation, and second, they commissioned a pre- and post-test evaluation from the outset.

The chapter then provides a detailed examination of an intervention that made use of a change model designed for change at the instructional core. It begins with

a detailed description of the design of the Catch-Up Programme. In the second section, I report on interview evidence from teachers who participated in its implementation. The final section presents the quantitative findings from the baseline and endline tests.

The Intersen Catch-Up Programme

The design of the lesson plans for the Catch-Up Programme was outsourced to an NGO but the selection of the learning materials, particularly the graded readers, was made by the leading literacy person on the provincial department team.

On the cover of the Term 1 Lesson Plans for the Intersen English Catch-Up Programme for Grades 6 & 7 is a line-drawing of a woman in running gear about to break the winning tape. On the tape are the words 'Fast Track Winner'. The illustration

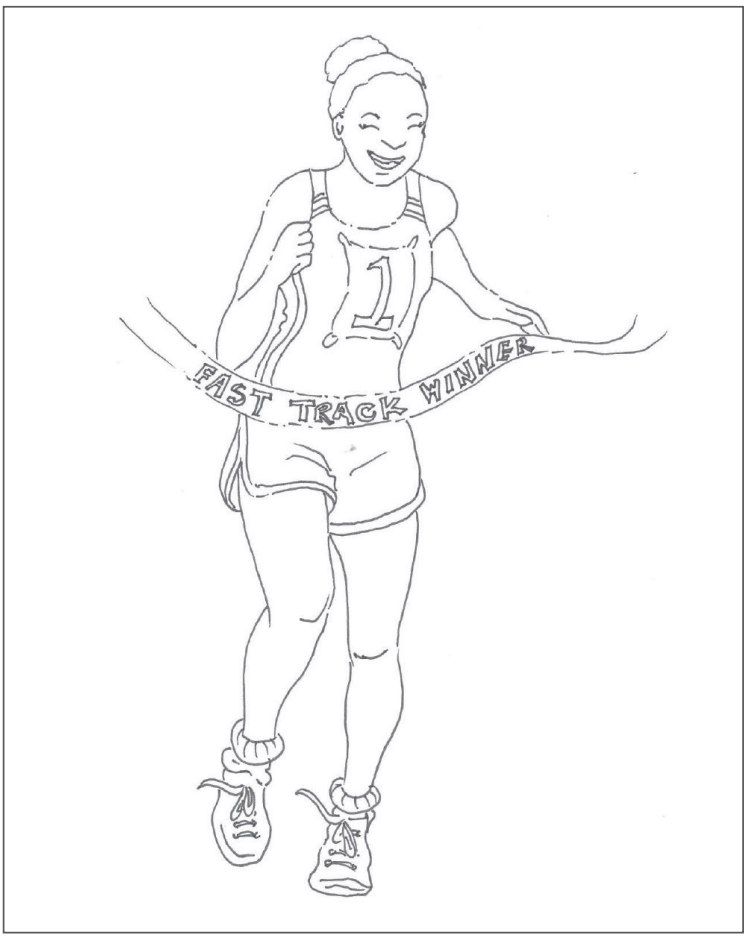


Figure 7.1: Picture on the cover of the Grades 6 and 7 Lesson Plan Guide

foreshadows a number of key features of these lesson plans, particularly the increased pace of teaching. The lesson plans would dramatically change the daily rhythms and routines of teaching, accelerating and intensifying the teachers' work rate in the English classes.

The first page of the lesson plan guide divided the first term into 11 weeks, with each week designated with a number—for example, Week 8—, and each numerical week was linked to a particular calendar week, for example, Week 8, Monday 5 March 2012 – Friday 9 March, 2012. Each calendar week had pre-specified assessment tasks. These seemingly simple weekly plans signalled to teachers that they would need to keep up, and that work assigned for the specific work week would have to be completed by the end of the calendar week so as to ensure that the learners were prepared for the assessment on the designated dates. The subsequent pages of the lesson plan guide spelled out in 11 brief 'notes', the template for the new programme, including information about the learning resources methodology, pacing and assessment.

The programme used six different learning resources for the classroom. The first and most obvious was the printed A4 black and white lesson plan guide itself.

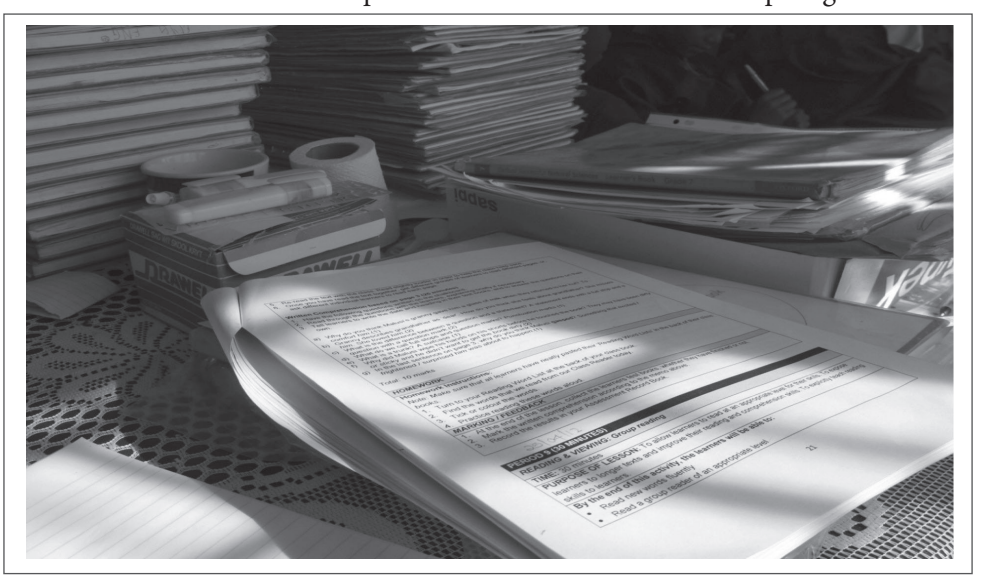


Figure 7.2: Lesson plans on a teacher's desk

The second was two A4 learner exercise books for each learner, one to write in during the regular class time and a second specifically for tests. The guide prescribed that the class exercise book was to be sent home every day and the test book at the end of the term. The four listening and teaching posters provided to each class covered four themes: In the Classroom, At the Zoo, On the Beach, and At the Hospital.

Table 7.1: Learning Resources for the Intersen Catch-Up Programme

Lesson Plan Programme Guide	1 per teacher
Exercise Books (Class and Test)	2 per learner
Listening and Speaking Posters	4 for the class
Graded Class Reading Books	12 titles, 20 copies per title per class
Reading Sheets	1 set per learner
Assessment Activities and Spelling Tests	1 per teacher
Assessment Record Book	1 per teacher

Source: Adapted from Gauteng Department of Education (2012)

The key learning resources provided to all Intersen classrooms were sets of 12 'reading' books. The lesson plan guide listed the book titles and the week that they were to be used in. Teachers were instructed to count the books at the end of every lesson to ensure that none had gone missing and to keep them stored 'neatly and carefully'. The selected titles were listed on the Gauteng Department of Education approved book list as Grade 2 and 3 books for English Home Language learners. The use of Foundation Phase readers for Grade 6 and 7 learners was based on the research (PIRLS, SACMEQ and ANA) that suggested that most learners in disadvantaged schools were three or more years behind the appropriate grade level in reading in English.

In addition to the A4 exercise books and the 240 reading books, the teachers received a set of 'reading sheets', one set per learner. The lesson plan guide instructed the teacher to ask their learners to paste these reading sheets into their class A4 exercise books. The reading sheets contained 'look and say' words that learners were expected to know the meaning of and commit to memory for the formal assessment. The 'look and say' words were derived from the reading books and constituted the core vocabulary and spelling words for the programme. The 'look and say' technique, however, did not dominate the Catch-Up Programme's systematic reading approach, but formed one of three distinct interconnected components along with a phonics programme and the readers, which teachers called the 'thin books'.

The final learning and teaching resource was a mark book, called 'the Assessment Record Book' according to the programme.

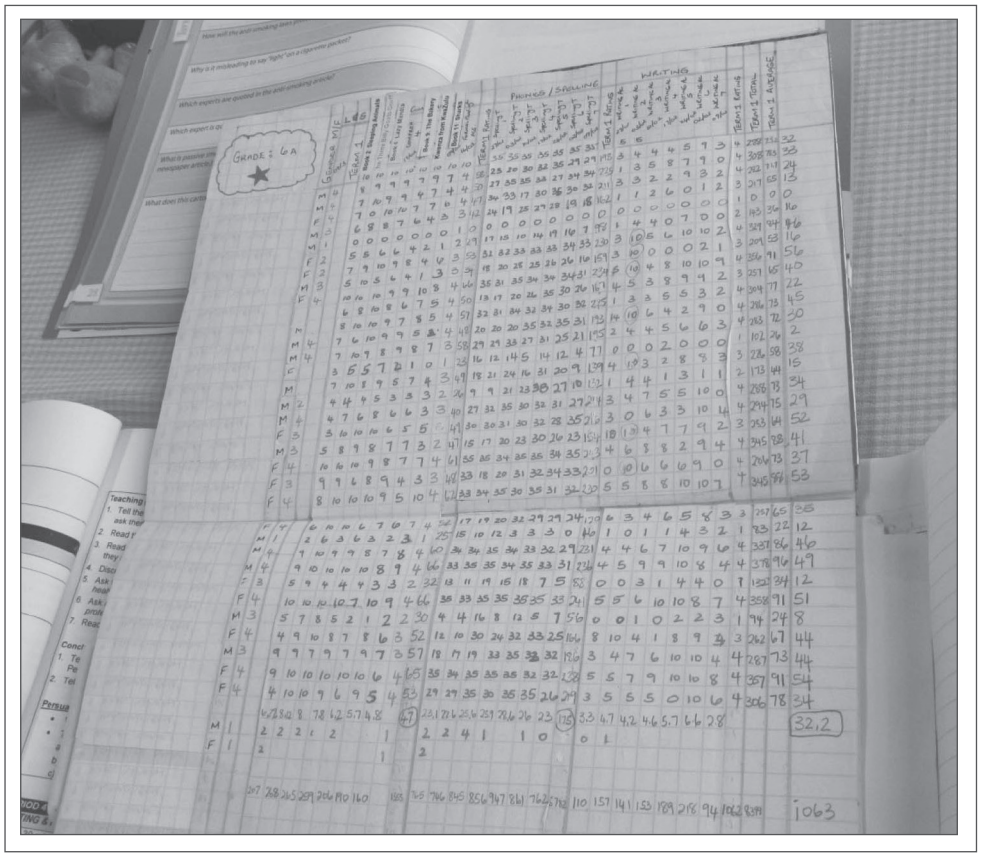


Figure 7.3: Assessment Record Book

The designers of the Catch-Up Programme prescribed a strict and consistent weekly teaching routine to be followed in the same sequence every week. The teaching week was divided into seven half-hour teaching periods. The teaching and the homework for each period was specified. Every week was to begin with a 'listening and speaking' task during which teachers would teach 10 sentences using the posters.[5] The second period was for phonics and spelling: two new sounds and related words as well as specific high-frequency words were introduced. The third period was devoted to teaching the 'look and say' words that would appear in the class reader for that week. During the fourth period, the teachers were expected to begin using the class reader assigned for the week. The tasks for the period

5 In the first two weeks of the term, teachers struggled to get through all 10 sentences. The teachers were encouraged to get through as many as they could, but would need to go on to the next task within the prescribed time period.

included reading aloud, shared reading and an oral comprehension exercise around the class reader. Period 5 was used for consolidation, and the sixth period for reading and writing. The final period of the week had two main activities, namely, writing and assessment. The assessment took the same form every week: a spelling test and a comprehension task. For each period, the guide specified the required homework. Save for the week during which there was to be formal assessment, each week would follow exactly the same format as the teacher worked systematically through the 12 graded class readers, the 4 posters, and 12 'look and say' word sheets.

The daily lesson plan guide provided a comprehensive description of each of the 70 lessons. A typical daily lesson plan began with a heading that specified the week number, day of the week and the date. The lesson time (number of minutes), lesson outcomes and lesson resources were all shown at the top of the page. The 30-minute lessons had either one or two activities. The bulk of the daily plans consisted of descriptions of these activities. The activities provided fairly detailed tasks per activity.

The lesson plan specified the questions that teachers had to write on the chalkboard and provided the answers (but told the teachers not to write these on the board). The 10 questions on the readers varied. Some were simple recall questions from the text (for example, name the fruits that they use to make the fruit salad); others required learners' own responses (which is your favourite?); and a few required slightly higher-order engagement (why do they add sugar over the fruit salad?).

The lesson period always ended with homework tasks. A typical homework assignment would have required learners to practise their 'look and say' words. The lesson plan guide instructed the teacher to tell the learners that they should memorise the words, making sure that they understood what they meant. Once they were certain they had mastered each word, the learners could put a tick against the words. Teachers were then guided to check a few learners per day orally at their desk.

The antithesis of the Curriculum 2005 outcomes-based approach, namely, scripted lesson plans, re-emerged in South African curriculum policy with the Foundations for Learning (FFL) Campaign after 2006. While building on the first generation of Foundations for Learning lesson plans, the GPLMS approach had a number of distinct characteristics. First, unlike the FFL lesson plans, GPLMS lesson plans were not optional; implementation was mandatory. Second, the GPLMS lessons had specific dates on which each of the lessons were to be taught. Third, the architecture of the lesson plans was simpler and clearer. The GPLMS lesson plan outcomes were straightforward and contained no jargon or references to the official curriculum policy. The GPLMS lessons had fewer lesson activities per period, and these were mostly shorter. The fundamental difference, however, was that the GPLMS lesson plans assumed that every teacher would be provided with a full set of quality

- Hand out the reading books.
- Try to ensure that no more than 3 learners have to share a book.
- Look at the cover of the book — ask the learners what they think the book is about (eating, fruit, school, etc).
- Tell the learners that you are going to read the book together. Start by reading the title of the book together.
- Turn to page 1. Read aloud with the class. Read slightly louder in order to help the class keep pace.
- After every page, ask a question about what you have read, for example: Why do we wash our hands?
- Make sure that the learners understand the meaning of what they are reading.
- Once you have read the whole book together, get different groups of learners to read different pages.
- Finally, read the book once more, this time asking different individual learners to read each page.
- Next, move on to the oral comprehension activity.
- Have the questions written on the chalkboard, and use this as a teaching opportunity.
- Explain to children that questions are often asked in the same way, 'and we are going to practise answering them together.'
- This will mean that when the learners see these kinds of questions, they will know how to answer them.
- Use the 'Guide to Comprehension Questions' at the end of this file to help you to teach this effectively.
- Once you have explained the questions to the learners, get learners to answer the questions.
- Direct questions at different learners — ask them by name.
- Always give different learners a chance to answer.
- Make sure you include all the learners in the class over a period of time.
- Let learners know if their answers are correct or incorrect. If an answer is incorrect, tell the learner, and ask someone else to answer.
- If a learner provides the wrong kind of answer, explain why it is wrong and let the learner try again.
- If any learners are really having a problem with the book, call them to your classroom at break or after school to do some extra work on the reading with them. Use the flashcards to teach them 'look and say' words.

Source: Adapted from Gauteng Department of Education (2012)

classroom learning and teaching resources. This allowed the designer of the lesson plans to specifically refer to resources that would be in their classroom. This saved teachers from the complex task of mapping lessons against resources (presuming teachers had those resources).

The lesson plans and the high-quality learning and teaching resources were fundamental for instructional change at scale in this model. The other component, that is, the professional development and support in the form of just-in-time training (conducted by the coach) and the ongoing on-site instructional coaching, was pivotal in shifting habits and routines in daily teaching practice. The coaches played a number of roles in the programme. They provided training to teachers in small groups, and they visited classrooms to model teaching practice, to observe, support and encourage teachers as they worked on the lesson plans. All coaches were themselves trained by one of the NGOs in the use of prescriptive protocols for coaching practice. During the training of the coaches, they used a tightly scripted training manual. Coaches were provided with a detailed facilitator's guide and supporting documentation.

How did the lesson plans, prescribed learning resources and coaching change instructional practices? It was assumed that together they would disrupt and re-engineer two core elements of practice. First, the lesson plans and the coaching would change how time was understood and used. The first page of the lesson plan guide clearly linked particular lessons to specific calendar days, thus specifying the pace at which the learning programme was to unfold. The pace of teaching was to remain the same even if teachers were absent or the day was interrupted for any reason. The responsibility or burden shifted to the teacher to keep up or catch-up with the pre-specified time frames. Within the lesson, teachers needed to increase their stamina to keep pace with the relentless forward motion of activities within the lesson plans. The role of the coaches was to assist teachers, and once trust was created, to push them harder to remain on track and to keep up. Faster pace, particularly around management tasks, meant more time for learning, which in turn would lead to improved outcomes. The prescribed weekly lesson routine provided a defined structure to school and lesson time. The uniform routines and rhythms structured into those lesson plans, once internalised, would allow teachers to cope with the increased pace.

Second, the lesson plans and the learning resources, complemented by the work of the coaches, expanded the teachers' pedagogic techniques and classroom management repertoire. One of the consistent findings in the literature (Fleisch, 2008; Carnoy, Chisholm & Chilisa, 2012; and Taylor, Van der Berg & Mabogoane, 2013) is the narrow range of activities and tasks teachers tend to use. The Catch-Up Programme lesson plans mandated a range of instructional methods and techniques. These included vocabulary development using the wall chart, graded reading using self-contained single-theme readers, systematic phonics, 'look and say' word lists, and writing and comprehension strategies. While teachers may have made use of some or even all of the methods or techniques at one time or another, the lesson plans provided a systematic and integrated framework within which each method or technique was

to be deployed sequentially over time through the structured framework. Teachers were to experience how the learning tasks embedded in each lesson built on each other, and how the various methods and techniques, for example, phonics and class reading, reinforced the learning pathway. The lesson plans were also to provide tangible guidance to teachers on resource and classroom management.

Teachers' talk

What did the teachers think about an instructional programme that aimed to disrupt and ultimately transform their existing classroom practice? In June 2012, as part of the work with the GPLMS, I undertook what I referred to as a rapid review of a purposive sample of five primary schools in the programme. Two lessons were observed in each school: in most instances it was a Grade 4 mathematics lesson and a Grade 7 English lesson. As the observations were conducted over five days, I was able to see how five different teachers were making use of roughly the same set of lessons. In addition to the classroom observations, I recorded interviews with teachers, principals and coaches, and took photographs of educational aspects of classrooms. In the interviews, I focused on teachers' views and experiences of the GPLMS and specifically how each of the components, that is, the lesson plans, learning resources, training and coaching, were working.

As would be expected, teachers were reticent, even resistant, to the new and unfamiliar approach and took some time to adapt to it and ultimately to adopt it as a core part of their daily work. One of the coaches observed:

> at the first training session, about 60 per cent of teachers were kind of accepting, and the balance were very anti, they didn't want it at all. They did not see the necessity. That was a rather difficult training session. During the first term doing the Catch-Up programme, they became more familiar with the methodology, now I have no more resistant teachers. I think the main reason they are coming on board, they are seeing that the learners are benefiting.

But once the teachers began to make use of it, the structured nature of the programme emerged as a positive feature. This, and the fact that teachers received a full 'package' that is, lesson plans with all the classroom resources provided and a coach who could help them use them, shifted teachers' perception of the value of the programme.

Scripted lesson plans

Coaches and teachers had a great deal to say about the lesson plans. But specifically on how the lesson plans impacted on practice, a coach had these insights:

> The good feature is the regularity of the programme. Teachers know exactly what they are doing, the sequence and what it is they are looking for. The lesson

> plan tells them exactly what to do. In the Catch-Up Programme — begin with the children, ask them to clear their desks, ask them to sit up straight, and look at you. Things like that. Do we ever think of telling people to do that, but the lesson plan told them to do that. And it's now with most of them, it has become a habit. This allows teachers to get learners' attention before they begin teaching. In small subtle ways like this, their teaching practice is changing. And more visual, they are coming to see how it helps them, not just to chalk and talk, they are learning to use other techniques also. They are becoming more aware of how learners respond to different techniques. As teachers become more comfortable in their zone, they are able to notice different things.

She signalled the ways in which the lesson plans provided teachers with new skills, both classroom management and pedagogic techniques, to add to their teaching repertoire. The coach was suggesting that the 'system' often assumed that teachers had some basic classroom management skills, but many did not. The lesson plans did not make this assumption but actually specified a range of techniques that many of the teachers began to make use of.

The theme of the value of the structure and prescription was a frequently recurring theme in the teacher interviews. As one teacher commented:

> It makes it easier for us. It is more structured. We were spoiled in Term 1 in English … we had our sentences, words, strips, comprehension, and told on which date was the test. It was structured. They even had the mark allocation which made it easier for us. The prep was done for us. It is a scheme-of-work.
>
> Lesson plans are very clear. The teachers can open the lesson plan and they know exactly what they need to do. It spells out everything for them.
>
> You know you're doing prepositions, which day to write the test, it was structured. And the marking and the mark allocation, it was like a scheme-of-work.
>
> Another good feature is the regularity of the programme. Teachers know exactly where they are going, the sequence and what it is they are looking for. The lesson plan tells them exactly what to do.
>
> The good thing, the structure is such that you know exactly what to do. And it's already prepared for the teacher, you know what to prep for. The lesson plans daily are already given to you. My prep book is exactly like that of my children, what is in my book is what is in my children's book. When you have different classes, the prep programme is much easier as I know exactly where I am for each class.

The teachers openly acknowledged the incentive associated with the GPLMS programme was that it reduced their workload and relieved them of the stress of daily lesson planning.

> I don't have to prepare, I just need to simply do it for my learners, the activity is here, the assessment is here, I am just stuck with a lot of marking. I am learning to do it faster. It has improved my teaching and it is interesting too.
>
> The first lot was structured it was like day 1, day 2, day 3, at the end of the week you test these words, and it was nice. It makes it easy for us. We were spoiled in term 1 in English.
>
> We had too much paper work, so we are happy, everything is worked out for us in the new lesson plans.
>
> I am happy with the lesson plans; it cuts down on your time.

Integration of learning materials into the lesson plans

The interviewed teachers appreciated the fact that the programme integrated the lesson plans with learning resources.

> And they gave us the resources and that was very good. And reading was incorporated. Different books all the time, and they were encouraged to read. So the thin books made them realise that they could read books. We have gone through many of the books and they are very excited to read the next book.

> I like the programme because it is integrated and the workbook has pictures, learners can point and speak about pictures. And there is writing in the workbook.

While the lesson plans had aspects that were familiar to teachers, that is, they had the look and feel of the old 'scheme-of-work', the class readers or 'thin books' were new to most teachers. Teachers soon realised the value of the carefully selected class readers.

> I love the class readers; it is getting them to read it.

> I like the graded readers because the children read a number of books, different series of books for more information.

> The readers are good because the children who are not so good at reading can read the simple books, and the fast readers can read more advanced books. Like the book about sharks. For those that know how to read, they need more advanced book. It builds the children's confidence, they say 'ma'am we have read so many books'. The stories are nice, they are interesting and colourful. And good quality, but I really wish they could give us more.

> They sent 20 readers, we have to share. They love the books, they love the books. They say, Ma'am can't we do English, Ma'am can't we do English? The books

> are very very good, Shuter and Shooter. But we have to share the books, the four Grade 4s.
>
> The reading material is the best part of the programme. It would not work without the reading materials. The fact that I have a book means that LTSM is not an issue. You don't have to worry about ordering the material as it is all done for you. It has been selected, it is good and appropriate. It is the only area that you have a lot of flexibility. If we did not have GPLMS, we would not have the resources as our book orders are always late, end of March.

Some teachers recognised that while the books themselves were important, the key to their success lay in the way they were mediated by the lesson plans, as one coach observed:

> Class readers were well received; they are topical, colourful and bright. The lesson plans spell out specifically how to use the class readers. For example, before they would not look at the cover, title, author, etc. They would read with the learners, reading aloud. Otherwise, they would let the learners read and they would listen. They would read a few pages and then stop and ask questions about the story and the pictures. They are short, you could easily manage a book in a lesson. They would never have a problem reading a book. They would then break into groups which would read aloud. In the catch-up reading, they were not encouraged to read individually. Now the children are full of confidence. In the Catch-Up programme, two books a week, now two books every two weeks. First book done orally, the second in their books for assessment. We have retained that pattern.

Coaching

Teachers talked extensively about the coaching process, their interaction with the coaches and the emotional side of the teacher-coach relationship. For example, the Grade 4 teacher described her engagement as follows:

> The coach is really supportive to me and listens to my reasons. I had a problem that I didn't do paragraphing, it meant that I had to go back and find additional time to teach it. The coach helped me solve the problem, helped me catch up.

It was not uncommon to hear emotive language about coaching and the coaches.

> It is very encouraging and very helpful. I love to have a coach. They take you from where you are and you have time to ask questions when you don't understand. And they are familiar with the material.

> The coaches come and assist us, that is very good. We have never had this in education.

The coach was described as a supporter and as a listener. The skills development role was also highlighted. A teacher described the specific teaching role that the coach played as follows:

> The coaches helped with demonstration. I don't know the sounds very well in the catch-up programme. She was friendly enough, where you are lacking. I am having a problem with sounds. After the demonstration I knew what to do. It was very great. She was presenting to us at another school. Planning ahead with GPLMS, punctuality and take time, planning, and adhere to the programme, is what I have learned.

While the supportive and teaching role of the coach is a recurring theme, the role of coaches in accountability was equally and surprisingly positively described. A Grade 7 teacher described it this way:

> They come and they see that the programme is being implemented. And it is being done correctly. That is why it is good having them around.

Another teacher pointed to the theme of accountability by contrasting the role of the coach to that of the district facilitators:

> They come and check that it's done by the book. It helps you know that you are doing it correctly. And that you are doing what is expected of you. If it isn't, the coach will tell you. The district facilitators, they were here last week, they have stopped inspection, and you go on your merry way.

Not all teachers were as positive about the implicit accountability role played by coaches:

> It is seen as very intrusive, and teacher gets stressed. They are here as coaches not inspectors. They are judging you, and that is good, but it does give some teachers ulcers. The person is not only in the classroom but also in the staffroom. There is some resentment, and it is growing.

Whether intentionally, or by default, the coaching process involves building trust, developing emotional relationships, skills transfer and holding teachers to account. This combination of attributes is captured in the assessment of one of the coaches:

> Maybe it's me, I push them, I will not take no for an answer. I have an idea of what they can do, and I insist that they do it. In the beginning the emphasis was on being critical, now I feel like we are friends, the pushing is less, they are truly with me. There has been a change in mindset. They also like the fact that I am always there, I take calls immediately. They feel that they have this safety net. If anything is worrying them, I am there to explain and to guide.

While further research would be necessary, there is evidence from the interviews that the education triple cocktail worked as a 'nudge' (Thaler & Sunstein, 1999), and that there is positive reinforcement through the lesson plans, the materials and the coaching to influence changes in teachers' instructional practice.

Critical comments

Notwithstanding teachers' overall positive sentiments about the programme, teachers raised three key concerns. They almost all felt that the pace of the programme was too fast, that the quantity of marking was too high, and that the programme seemed to leave the slower learners behind.

On the pace, teachers commented:

> The lesson plans, the pace at which it moves is a bit too fast for our learners. It is really fast paced . . .

> Teachers struggle with the pace of the programme.

> But the teachers are 'crying' about too many tasks. They cannot finish the weekly work plan.

Teachers said the following about the work that needed to be done at home:

> Last term my boot was full every day. I had these piles of marking.

> There is a lot of marking in the first term, but I am happy for the second term. It's a big impact, big impact honestly. But the marking, wow.

> Teachers feel that they are spending their whole lives marking. How are we going to address the problem of marking, we are going to have teacher burn-out.

> The marking was a bit much in the first term, but they have reduced it in the second term. So it is much better.

> It is good to a certain extent, but the workload for teachers is quite a bit.

On differentiation, teachers commented:

> There is no room for the slow learners. We end up concentrating on the fast learners. We don't think it is fair for the learners who are struggling.

> GPLMS was fast tracking the learners. But the kids that are slow are becoming slower. We need a special class, and a remedial teacher to deal with the slow learners.

Only one teacher complained that the lesson plans undermined her professional autonomy:

> You know that the programme is given, the prep is given, you are expected to follow. You know it limits your creativity. I feel a little bit bottled.

Teachers described how the Catch-Up Programme accelerated teaching. What made this palatable were three features of the programme. First, the standardised lesson plans freed them from the onerous task of developing daily lesson plans. Second, because all the materials were purchased and delivered to classrooms, they did not have to worry about getting the learning resources. And third, they could see measurable improvements in learners' reading and learners' attitudes towards reading. The learning materials were without doubt one of the big attractions of the programme. Teachers had seldom, if ever, had a sufficient quantity of high-quality learning materials to work with. The coaches, however, played a critical role. The coaches created a trusting environment within which they could take the new practice forward. They provided on-site training not only in the use of the lesson plans and class books, but on a range of other related classroom management topics. And for some, the coaches put in place a friendly, unthreatening form of accountability.

Evaluation

The teachers' discourse provides powerful insights into how and why the prescriptive approach to large-scale instructional change may work for teachers. What the teachers' discourse does not do and cannot do is provide us with evidence of its impact on measurable learner performance. To merit implementing the Catch-Up Programme and this particular model of large-scale instructional change system-

wide beyond Gauteng, it would be necessary not only to measure the difference between learner performance before and after the intervention, but to have robust counterfactual evidence.[6]

The GPLMS managers were determined to gather information on learner performance, with a specific eye to assessing the extent to which the Catch-Up Programme impacted four distinct skills, that is, spelling, language, comprehension and writing. In addition, the GPLMS managers were also keen to assess the relative effectiveness of the various NGOs, and different improvement patterns in the districts.

Two assessment tools were developed, one for learners in Grades 4 and 5, and a second for learners in Grades 6 and 7. Even though many Intermediate Phase teachers taught multiple classes, to contain costs, the tests were restricted to one class per teacher. The teachers themselves administered the test and recorded the results, once at the beginning of the first term of 2012 (pre-test), and again 13 weeks later once the Catch-Up Programme had been completed. A database was created on which scores of each learner were captured. In addition to learner performance, the database contained linked information about teachers, coaches, NGOs and districts. All class data was matched — pre-test and post-test — using key identifiers to ensure that the same classes were being measured to allow for comparisons to be made. The final sample consisted of 1 570 classes, including 45 per cent of English Intermediate Phase teachers, in the project.

The pre-test results were not unexpected (Figure 7.4). While the average score for the sample was 37 per cent, the skill that learners did least well on was writing. This is consistent with the findings from the SACMEQ and PIRLS and the National School Effectiveness Study. However, the average score of the sample increased by 15 percentage points, with gains of roughly the same magnitude on all four skills. The post-test showed that as a result of the Catch-Up Programme, learners' spelling in particular improved significantly.

When the analysis of average scores was done across the grades, the pre-test results showed a steady upward progression (Figure 7.5). Again it could have been predicted, particularly the gains between Grades 4 and 5 and Grades 6 and 7, as these learners took the same tests. As with the case of gains by skill, although the learners' pre-test scores showed marked different starting points, the post-test showed gains of roughly the same magnitude (between 14 per cent and 16 per cent) across the four grades.

6 In the next chapter, I describe such a study of the Catch-Up Programme. Although randomised control trials (RCTs) are thought to provide the most valid impact evaluation results, there are a range of logistical problems associated with this method. Given the absence of random assignment before the intervention, alternative methods could be used to provide counterfactual evidence of the effectiveness of the intervention and approach. Counterfactuals could be created through 'matching', ie finding schools for the control group that match as closely as possible participants of the treatment group on key socio-economic and demographic characteristics.

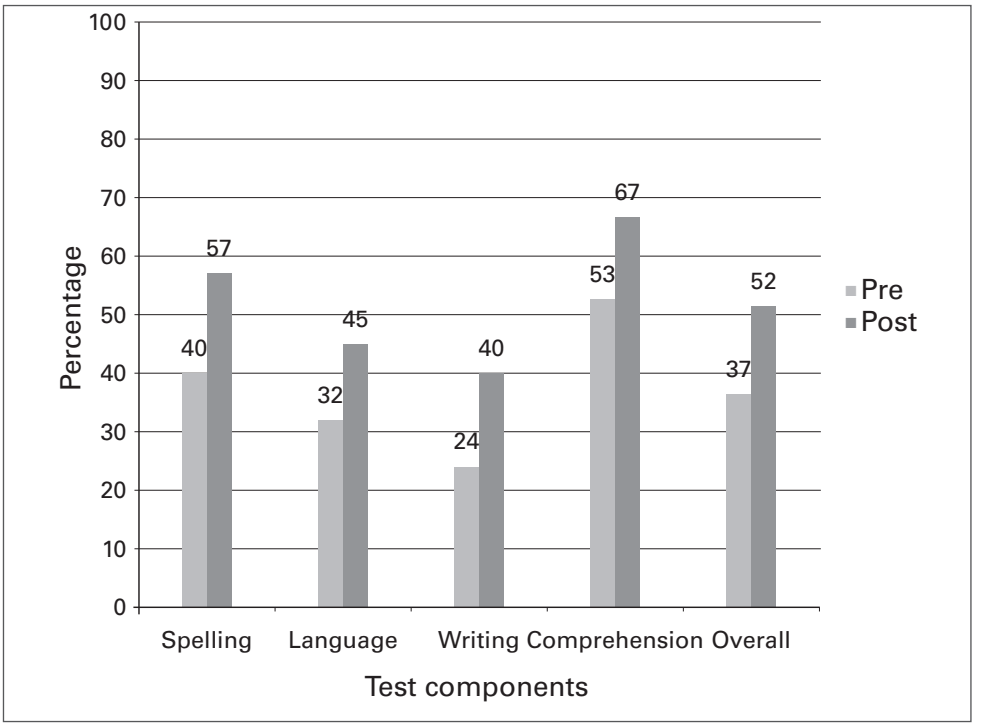


Figure 7.4: Results by literacy component
Source: Adapted from Hellman (2012)

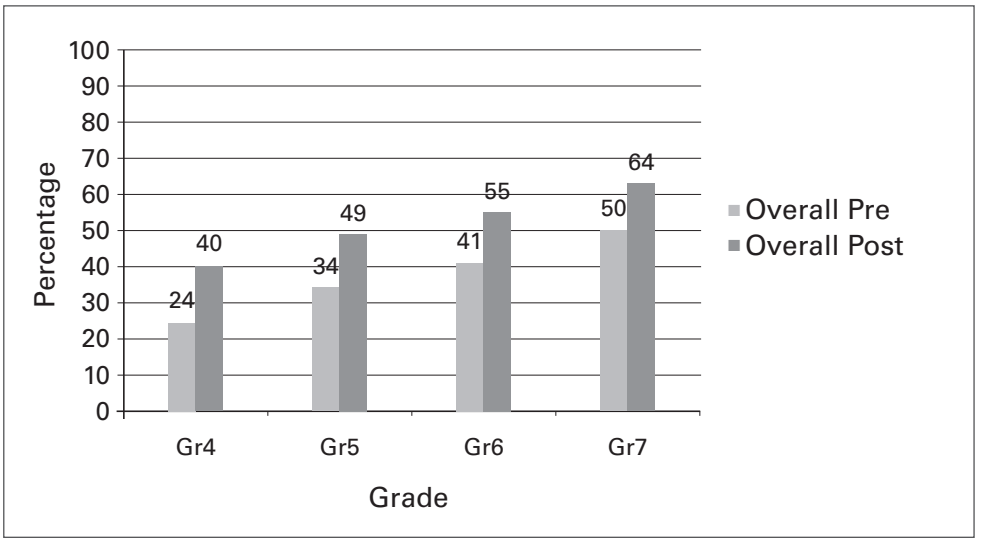


Figure 7.5: Results by grade
Source: Adapted from Hellman (2012)

When the database was disaggregated by NGOs, the between-NGO variance on the pre-test was not large (Figure 7.6). But again, as with the skills and the grade gains, the NGOs recorded similar gains. This pattern is also evident in the disaggregated analysis of district gains (Figure 7.7).

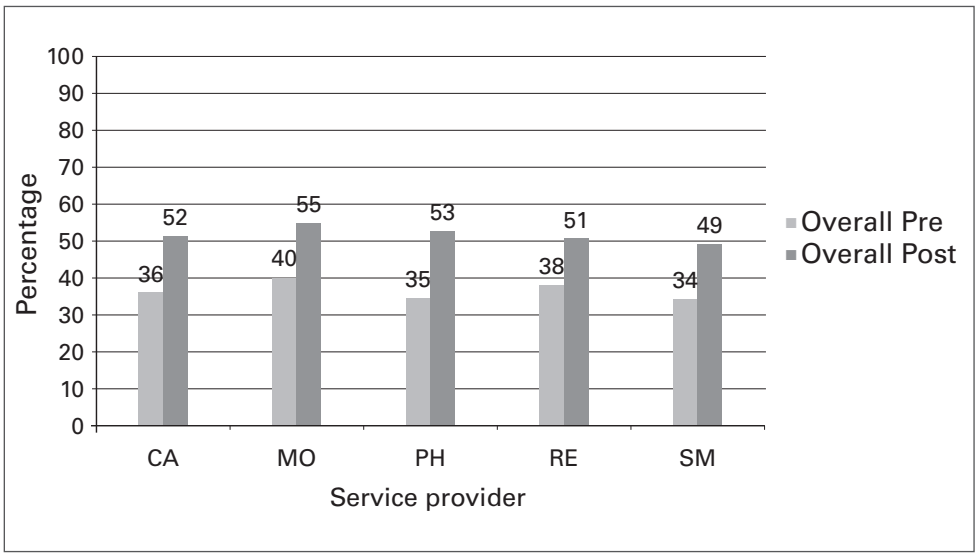


Figure 7.6: Results by service provider
Source: Adapted from Hellman (2012)

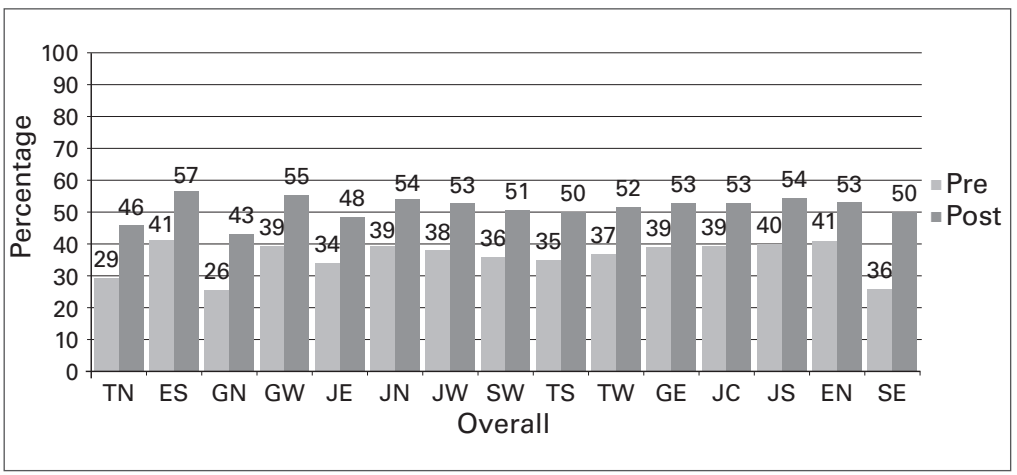


Figure 7.7: Results by district
Source: Adapted from Hellman (2012)

These figures suggest that while not all learners were on the same level of achievement at the start of the intervention, across skills, NGOs and districts, the magnitude of gains made by learners was roughly of the same order. The Catch-Up Programme seemed to be working equally well with all NGOs, and improving all First Additional Language English skills. Overall, the Programme's striking characteristic is that irrespective of grade, NGO and district, it seems to be having an exceptionally consistent effect.

While the average scores provide important findings about the magnitude of improvement for the sample, an analysis of the distribution patterns provides additional insights about the efficiency of the programme. Following on the categorisation in the Annual National Assessment, the evaluators scored learners according to four categories or levels. Learners were placed on Level 1 if they scored below 25 per cent; Level 2, 26–50 per cent; Level 3, 51–75 per cent; and above 76 per cent on Level 4. (These are not identical to the four ANA levels.) The study then analysed the percentage of the sample on each of the levels for the pre-test and post-test. This analysis revealed a noteworthy trend. As a result of the Programme, the percentage of very weak learners in the sample (those that scored below 25 per cent) declined from a quarter of the sample to 4 per cent. At the other end of the spectrum, only 1 out of 100 learners who did the pre-test scored above 76. This percentage increased to 6 per cent. A portion of learners appear to have improved sufficiently to move them from Level 2 (26–50 per cent) to Level 3 (51–75 per cent).

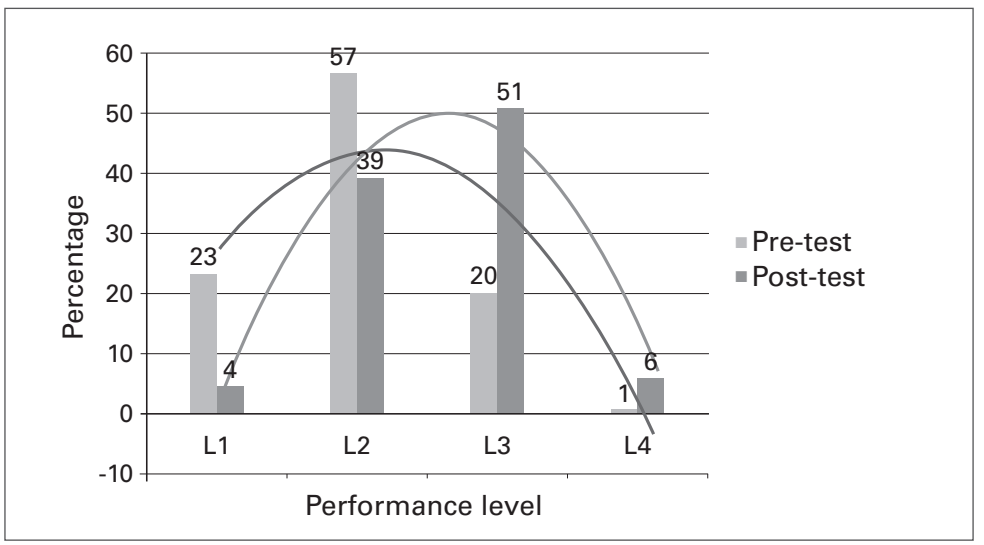


Figure 7.8: Distribution across the four levels, overall
Source: Adapted from Hellman (2012)

While the sample average increased by only 15 percentage points, the combined percentage of learners on Level 3 and Level 4 increased dramatically from 21 per cent to 56 per cent.

Conclusion

Research from the GPLMS, the first effort to implement the education triple cocktail model, shows a number of powerful insights about instructional change in South African primary schools. The vast majority of primary school children do not get to do enough learning tasks in the classroom every day and across the year; the learning tasks that they do are not in the main pedagogically sound; and these tasks are not generally part of an enacted systematic curriculum. The lesson plans, learning resources and personalised coaching address many of these problems but, if and only if, teachers teach using the programme. The lesson plans have the potential to substantially increase the time for learning, through the lesson/date scheduling and better classroom management.

Possibly more importantly, the pace of the lessons has speeded up, teaching is moving faster, and more content is covered in a single lesson. The scripted programme and structured approach also ensure a better balance of tasks, tasks that make use of varied and complementary methods and techniques: an optimal mix of listen-and-speak tasks that enhance vocabulary the learning of high-frequency words from the word lists, tasks that encourage reading fluency and reading enjoyment with the 'thin books' being read out aloud as a class, learning to use phonics to decode unfamiliar words and tasks that teach extended writing. And while the tasks themselves are crucial, it is equally important that they are experienced as a coherent whole that is sequenced to ensure optimal development. The store of knowledge is ever increasing, not too fast but not too slow. The number of words that have been mastered increases, enabling great fluency in reading and writing.

Possibly one of the most important lessons for changing instructional practice is to start from where teachers and learners are. This is the mistake that South Africa consistently makes. No matter how carefully the tasks are planned, aligned, sequenced and balanced, if the gap between the schoolchildren's current knowledge and skills and the target knowledge and skills is too wide, the entire enterprise is bound to end in tears.

Referencing Elmore, the teachers, as part of a system, can only learn the new practice by doing the new practice, and having the new practice embedded in the lesson plans, the books and the advice of their coaches. If they do not use the lesson plans, the books and the advice of their coaches, or do what Berman and McLaughlin called 'mutual adaptation' (1974), taking whatever can be taken without fundamentally altering the existing pattern of teaching, then the process will be of little value. To learn to do the work by doing the work means actually doing what

is really intended with fidelity. After a time, the rhythms and routines become the new normal. And with all the right support that is, school management, district capacity and political leadership, the thousands of new classroom tasks begin to knit together to form the new practice.

The Catch-Up Programme case study seemed to show that the education triple-cocktail model works and provides some insights into possible mechanisms. But would it really stand up in the face of studies that make use of more rigorous methods? The next chapter provides the answer.

Chapter 8
Compelling evidence

Introduction

This book addresses one of those transformational changes that take place in education systems. At present, South Africa effectively has two education systems under one set of legislation: a 'middle-class system' serving formerly white towns and city suburbs of South Africa; and a 'working-class system' serving black children living in deep rural areas, urban townships and informal settlements. The middle-class system has largely been inherited from the well-resourced remnants of the apartheid education departments, with elaborate physical infrastructure and strong instructional practices. As a consequence, in this system, the majority of its learners (but not all) become successful learners (eg proficient readers) by the end of the early grades and most go on to complete the full cycle of 12 years of formal schooling on time and get awarded a senior school certificate. The other system, which has roots that go back further than Bantu Education, teaches the majority of its learners only the rudiments of literacy and numeracy in the first few years of schooling, and less than half successfully complete the full school cycle. This working-class system as a 'system' with all its interconnected components is simply not geared to teaching the majority of children to be academically proficient. What it does is inculcate institutional habits and it works as a screening mechanism, filtering out a small group of outlier children (the 'clevers'). It is this select group of children who will go on to enjoy the benefits of post-secondary education.

Although achieving equity within the middle-class system remains a priority, the deep change challenge in South African education is to transform the working-class system so that the majority of school children achieve the official curriculum's stated learning outcomes. To put it in the context of one of the debates about educational change, the change that is required cannot be characterised as either 'innovation' or 'improvement'. The notion of educational change as innovation is associated with the avant-garde, novel or unique. While effective literacy instruction is not common in working-class schools, there is very little that is revolutionary or even innovative about the reading-teaching practices that need to emerge in working-class schools. The change that is required is also not 'improvement', which assumes gradual and incremental betterment. What we are talking about is a radical shift to be undertaken in a relatively short time frame (less than 10 years)

and which is to occur across an entire nation in thousands of schools and tens of thousands of classrooms.

The transition of the working-class system is both about more resources to the early grades and about teaching in this 'system'. This book has examined the latter. While resources are critical, the existing patterns, routines and rhythms associated with early grade instruction in working-class schools are maintained both by institutional school cultures and interlocking provincial and national instructional policies.

The focus of this book has been on understanding what it would take to fracture residual instructional practices and replace them with practices that work better and can be institutionalised across an entire system. We know that mostly change initiatives have failed either because they did not get to the instructional core, or they did not speak to the wide policy environment; and in most instances it was both. Given this, what kind of initiative and models of change would be powerful enough to disrupt existing instructional patterns and help teachers replace them with better practices that will endure? In earlier chapters I have described the features and characteristics of such initiatives, and the model of change that they embody. The aim of this chapter is to examine the rigorous evidence for it.

In my view, this evidence needs to be convincing and conclusive. It also needs to be rigorous and robust. 'Rigour' assumes the thorough and systematic use of the best social science methods available. 'Robust' means both getting similar results with different statistical models within a study and ensuring that the findings are replicable, across settings and school subjects.

Table 8.1: Studies in the South African Research Early Grade Learning Programme

	Name of study	Acronym	Research method	Time-frame	Focus	School size
Early Grade Learning	Gauteng Primary Literacy and Mathematics Strategy	GPLMS	Quasi-RDD, natural experiment	2011–2014	Primary school literacy and mathematics	1040 Schools
	Reading Catch-Up Study	RCUP	RCT	2014	Gr 4 10-week remedial programme in English	100 schools
	Early Grade Reading Study I & extension	EGRS I	RCT, qualitative case studies	2015–2016	Gr 1–2 Setswana literacy	230 schools
	Early Grade Reading Study I Expansion	EGRS I	RCT	2017	Gr 3 Setswana literacy	230 schools
	Early Grade Reading Study II	EGRS II	RCT	2016–2017	Gr 1–2 English literacy	180 schools

This chapter brings together the findings of a series of research studies that were designed to test the efficacy of the intervention model that aimed to improve instructional practice and learning outcomes at scale, particularly around early grade reading. The specific model, the education triple cocktail, is the combination structured programme that interlinks the provision of prescriptive lesson plans, learning and learner resources, and instructional coaching. The purpose of presenting the findings on the education triple cocktail from different studies, and from different research sites at differing time points is to contribute to building a rigorous knowledge base on change at the instructional core.

This current research programme began in 2011 with efforts to evaluate Gauteng's system-wide initiative (GPLMS) to improve primary school learning. When it became clear that to do different waves of large-scale classroom observations would be too expensive, the provincial department chose to focus on the findings from existing learner datasets. Building on the findings of the first Regression Discontinuity Design study/Natural Experiment of Gauteng Primary Language and Mathematics Strategy (Fleisch & Schöer, 2014; Fleisch et al, 2016), the system-wide instructional reform research programme took a major step forward with the implementation of the Randomised Control Trial (RCT) replication study that was designed to replicate an earlier study, which showed a strong impact of a shorter version of the triple-cocktail model (Fleisch et al, 2017a; Fleisch et al, 2017b).

The heart of the chapter, however, is the summary of the findings of the Early Grade Reading Study, that is, South Africa's largest and most ambitious randomised control trial that aimed to test not just the impact but cost-effectiveness of various models of instructional reform (Cilliers et al, 2016; Taylor, 2017).

What the international literature tells us

The focus of this chapter is on the findings of the South African research programme concerning the impact of structured pedagogic interventions on instructional practice and subsequent learning outcomes. It would, however, be myopic if I failed to locate this research in the field internationally. The increase in the number of similar studies, and the rise in meta-analyses and systematic reviews of multiple studies of this kind is an indicator of the intense global interest in large-scale instructional reform.

The first of the current-generation systematic reviews of studies of the impact of initiatives on learner academic performance was Patrick McEwan's (2015) systematic review of randomised control trials of intervention initiatives in sub-Saharan Africa. In an earlier version of his review, McEwan found that combined programmes, initiatives that used more than one type of resource or service, were most effective. This analysis of dozens of studies showed that teacher training, in particular, when combined with other types of intervention, had the greatest impact on learning outcomes. Possibly the best known of these multiple, study reviews was conducted

by the International Initiative for Impact Evaluation, or 3ie. Reviewing hundreds of studies done in many parts of the Global South, Snilstveit and colleagues (2016) found that different kinds of initiatives work for distinct educational problems. For example, they found that cash transfers (providing small cash incentives to parents) was the most effective way to increase school enrolment, reduce dropout and ensure that children stay in schools to completion. However, these financial incentives to parents did not improve learning outcomes.

What works to improve learning are structured pedagogic programmes. These are initiatives that provide customised curricula (like lesson plans), give teachers new instructional approaches and the training to use them, and make available new and better educational materials for their learners to use in their classrooms. Glewwe and Muralidharan (2015), in their conceptual review of a number of research papers, likewise found that interventions that focused on improved pedagogy (especially supplementary instruction to children lagging behind) were particularly effective. But they also point to interventions that improve school governance and teacher accountability. Two other independent researchers, Masino and Niño-Zarazúa (2016), came to somewhat different conclusions: their review suggested that initiatives that combine accountability measures with supply-side interventions were most impactful.

Although systematic reviewers claim to summarise only rigorously designed studies (those with strong counterfactuals), their conclusions are not always the same. Two World Bank researchers, Evans and Popova (2016), have recently attempted to deal with these discrepancies. After careful scrutiny of half a dozen systematic reviews they found some basic common ground:

> across the reviews, two classes of programs are recommended with some consistency. Pedagogical interventions that tailor teaching to the student learning levels — either teacher-led or facilitated by adaptive learning software — are effective at improving student test scores, as are individualized, repeated teacher training interventions often associated with a specific task or tool. (Evans & Popova, 2016:242)

Evans and Popova's (2016) analysis has not been without critics. Haddaway, Land and Macura (2016) accuse them and the authors of other systematic reviews of using methods that are at wide variance with standards commonly used in scientific systematic reviews. Other critics focus not so much on the systematic review approach but on the experimental methods used in the underlying studies. For example, Angus Deaton (economics Nobel Prize winner) and Nancy Cartwright (2016) argue that experimental trials seldom discover the actual 'mechanisms' that make the interventions work. Without identifying the mechanisms, Deaton and Cartwright argue, it is almost impossible to assess what can and cannot be transferred or scaled up. Nadel and Pritchett's (2016) critique is more or less in line with Haddaway, Land

and Macura (2016). They argue that including different interventions that appear to be in the same 'class' or 'type' is misleading. The obvious example would be comparing the GPLMS/EGRS model (South Africa) to the Tusome model (Kenya). Both appear to be structured pedagogic programmes that combine different components, but the GPLMS's core training is two days twice a year and monthly coaching by an external professional, whereas Tusome provides seven days of training and more regular coaching but from county education officials.

GPLMS

The absence of robust evaluation of the effectiveness of intervention initiatives was the catalyst for the systematic study of models that promise to fast-track improvement in early learning outcomes across entire systems. One of the more promising initiatives was the Gauteng Primary Language and Mathematics Strategy (GPLMS) (2011–2014), a provincial programme that made use of a three-component intervention model to improve outcomes in literacy (in African languages and English) and mathematics. It has served as the springboard for both research and programme development over the past few years.

My involvement with the Gauteng Primary Language and Mathematics Strategy was not simply as a researcher. I played a central role in developing the original intervention design and was involved as an adviser in the first three years of the provincial intervention. In that period I spent a considerable amount of time in GPLMS classrooms, observing teachers using the lesson plans, talking to teachers about their relationship with the coaches and new education materials. As I showed in the last chapter, many teachers liked the programme but complained that the intervention meant teaching got a whole lot more intense. I had a 'gut' sense that classroom practices were beginning to shift in real ways. That being said, the provincial team was mindful that rigorous evidence was needed. Although a budget was set aside for a formal independent evaluation, a decision was taken during the first year of implementation that the Annual National Assessment data would be used to assess the impact that the programme was having on the original 793 and later expanded group of 1 040 schools.

Figures 8.1 and 8.2 show the original test distribution of GPLMS schools and non-GPLMS schools on the 2008 Systemic Evaluation, as well as the subsequent performances of the two groups of schools in the 2011 and 2012 ANAs.

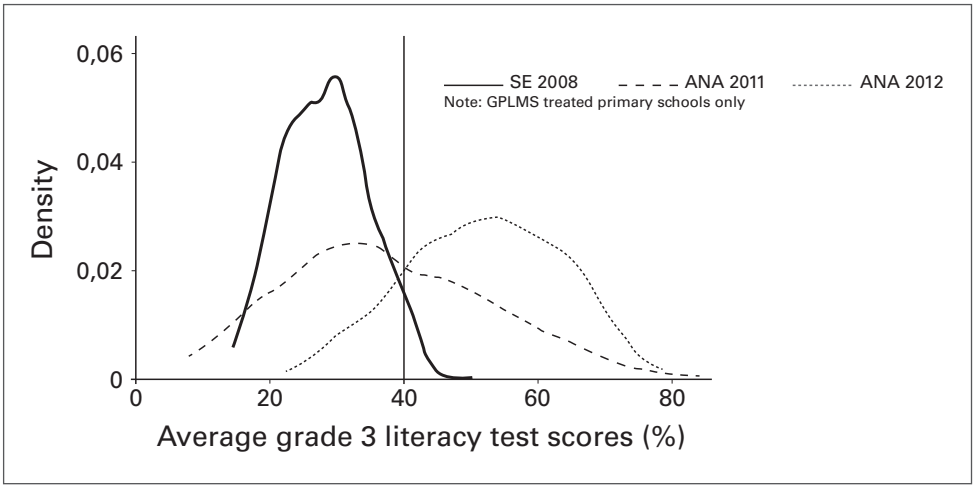


Figure 8.1: Distribution of the Grade 3 literacy scores in the 2008 Systemic Evaluation, 2011 ANA and 2012 ANA, GPLMS

Note: While the threshold to be designated as a GPLMS school was set at or below 40 per cent in the literacy section of the Systematic Evaluation, a small number of primary schools that actually performed below the cut-off were never treated and a small number of schools were.

Source: Fleisch & Schöer (2014)

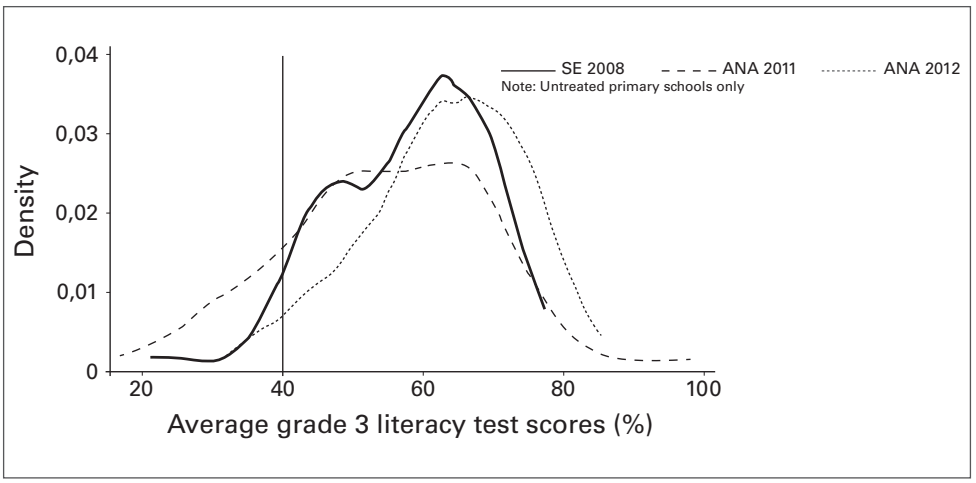


Figure 8.2: Distribution of the Grade 3 literacy scores in the 2008 Systemic Evaluation, 2011 ANA and 2012 ANA, Non-GPLMS

Source: Fleisch & Schöer (2014)

The frequency distribution in Figure 8.1 showed the initial performance of schools that were assigned to the GPLMS based on the 2008 systemic evaluaton scores, and how the majority of the GPLMS schools still underperformed in the 2011 ANA. However, by 2012, the same sample of previously underperforming schools dramatically improved their performance. These appeared to be major gains. This was most evident when we compared the distribution of schools in the programme to schools outside the programme. The frequency distribution for non-GPLMS schools in Figure 8.2 reflects much less change between the different test instruments in 2008, 2011, and 2012. When comparing the performance of GPLMS schools and non-GPLMS schools over the three test instruments, GPLMS schools show a great improvement relative to the non-GPLMS schools.

But are these statistics credible? Volker Schöer and I make the point that the changes in the distribution of test averages and the overall distributions may be driven by a number of factors, some unrelated to the intervention. It is likely that the design and implementation of ANA tests themselves might have allowed weaker learners to perform relatively well. If this is the case, the dramatic improvement in GPLMS distribution between 2008 and 2012 might not be an indicator of 'improvement for real', but simply a test instrument effect (Fleisch & Schöer, 2014).

A second effort to get a more precise estimate of the impact that the GPLMS intervention was having and control for a possible instrument effect, was to forgo comparison with earlier tests and construct a research design that was as close to a perfect counterfactual design as possible. Essentially, this required us to compare two groups of schools with very similar demographic characteristics, the main difference being that the one group got the GPLMS intervention and the other did not. Using the 2013 ANA results, we did such an analysis using a hybrid natural experiment/regression discontinuity design (RDD). We compared two groups of schools that clustered immediately above or below 36–44 per cent — the original 2008 systemic evaluation cut-off line of 40 per cent and schools that were incorrectly assigned either to the GPLMS or not-GPLMS group (Fleisch & Schöer, 2014). In 2008, the two groups of schools had the same average performance, but by 2013, the GPLMS schools were a full 10 per cent points higher than the comparison group of schools.

The natural experiment/RDD study, particularly the evidence from the sample of schools around the assignment threshold, and the almost 10 per cent point difference between the control and the intervention groups, provided convincing but not conclusive evidence that the GPLMS model can effect instructional practice at scale and by extension raise learner outcomes at scale. The key limitation of this study was that the impact was measured only with a very specific sub-population within the overall distribution of schools, that is, the academically stronger schools within the underperforming group.

This begs the question about what impact the intervention might be having for

the systems as a whole. While certainly not as rigorous as our study that used the ANA 2013 results, the publication of the ANA 2014 results did provide some clues.

Table 8.2: Average scores in numeracy test sections for GPLMS and non-GPLMS schools across years and grades

	Sample of schools around assignment threshold (36–44 per cent)			
Grade 3	Non-GPLMS	GPLMS	Difference	Standard deviation
2008	45.1	44.5	-0.6	18.8
2011	36.7	36.5	-0.2	15.6
2012	42.2	49.1	6.9**	14.4
2013	51.2	61.0	9.7***	14.1

Notes: *** p<0.01, ** p<0.05, * p<0.1
Source: Fleisch et al (2016)

Table 8.3: Percentage of Learners ≥50 per cent on the Verification Annual National Assessment, 2014

	Gr 3		**Gr 6**		
	Math	**Lang**	**Math**	**HL**	**FAL**
Eastern Cape	52.3	51.0	22.0	53.7	24.2
Free State	65.2	64.8	41.0	69.3	45.3
Gauteng	73.6	70.5	53.1	90.7	58.2
KwaZulu-Natal	56.4	56.7	31.4	73.3	32.1
Limpopo	33.6	42.5	15.9	54.6	24.0
Mpumalanga	54.9	56.6	27.1	83.8	34.9
Northern Cape	48.9	44.8	24.4	63.1	28.8
North West	39.4	51.2	20.0	74.5	31.6
Western Cape	68.0	64.2	44.9	79.9	52.2
National	55.6	56.5	32.4	74.6	36.1
Diff GP and next highest	*+5.6*	*+5.7*	*+8.2*	*+6.9*	*+6.0*

Source: Department of Education (2014)

At its peak in the middle of 2014, the GPLMS was being implemented in 1 040 underperforming primary schools in Grades 1 to 7 in Home Language, English and English as First Additional Language and mathematics. Using the indicator, the percentage of learners in the verification schools who scored at or above 50 per cent on the ANA 2014 tests, Gauteng was outperforming all other provinces on all tests by a substantial margin (Table 8.3).

The Pinetown study

The three sources of GPLMS evidence that speak to the efficacy of the triple-cocktail model of system-wide instructional reform are all suggestive or even encouraging, but even put together they certainly are not conclusive. Conclusive evidence would require the use of the most rigorous methods to measure impact, specifically the use of the so-called gold standard, the randomised control trial, and replication studies (the latter would be necessary to address the problem of external validity, that is, the extent to which the results could be generalised to other systems). The first explicit research attempt to apply the randomised control trial method to the education triple-cocktail model was undertaken in Pinetown, KwaZulu-Natal in 2014. As described in Chapter 7, the original GPLMS intervention had developed a short 11 week catch-up or remedial programme for Grade 4 to 7 learners who had not benefitted from improved early grade teaching. The remedial programme used the same basic model of change, the combination of prescriptive lesson plans, good-quality learning books, including appropriately levelled books, and instructional coaches. Unlike the main intervention, the official CAPS curriculum was suspended for the 11-weeks so that the entire grade cohort had an opportunity to upgrade their basic reading skills.

The Pinetown study was geared to answer the question, 'By how much would the intervention change the learners' performances compared to how they would have performed if they had not been part of the intervention?' To answer this kind of question requires an intervention and a counterfactual (a close to identical group). Assuming that the intervention group and the almost identical group are more or less on par at the start, then the difference at the endpoint would be the impact of the intervention. (For a full discussion of the methodology, see Fleisch et al, 2017b.)

The best way to establish a counterfactual is to assign a group of learners to an intervention (or treatment) group and a comparison (or 'control') group using a lottery. This 'random assignment' is by chance and not by any specific characteristic of learners in the group; there would be no reason to expect the treatment group to be different from the control group. According to Fleisch et al (2017b):

> Since receiving the programme would be completely random, the two groups should be similar in all observable characteristics. But even more powerfully, if

no individual could choose to be in either group, the two groups should also be similar in all un-observable characteristics. It is for these reasons that the RCT design is considered the most reliable way of constructing a counterfactual.

Big enough samples will ensure that the two groups are going to have an equivalent number of different kinds of school, that is, low and relatively high performers. Using the RCT method with two key assumptions, that is, schools were assigned to the treatment and control groups on a genuinely random basis, and we had an adequate sample size to make sufficiently precise inferences to the wider school population, the Pinetown study was in a position to make rigorous claims about change at the instructional core.

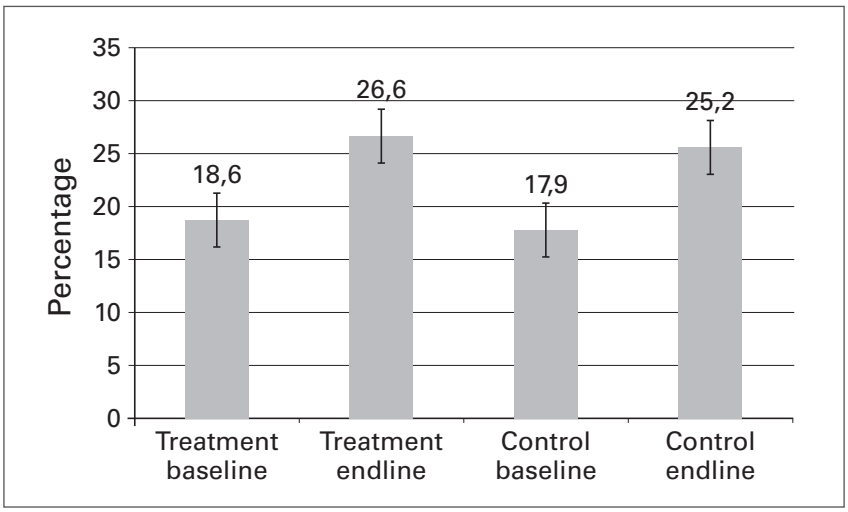


Figure 8.3: RCUP baseline and endline results for treatment and control schools, 2014
Source: Fleisch et al (2017a)

Although there was a substantial improvement in test scores between the pre- and the post-test, the average performance on the postline test was very similar for the treatment and control groups. This shows that on average for the groups, there was no added value in being part of the intervention. In simple terms, the large-scale, term-long remedial intervention that used the education triple-cocktail model of change did not seem to work in the Pinetown District of the KwaZulu-Natal Department of Education.

Why would there be such a dramatic rise in test scores in the control group? It may simply be that soon after beginning instruction in English as the language of instruction (as occurs in Grade 4), the schoolchildren would rapidly gain new English vocabulary. Another possibility relates to the Hawthorne effect, namely that irrespective of whether

a school was assigned to the control group or the treatment group, all the schools were subject to external scrutiny, particularly performance testing (that is, pre- and post-testing). A third explanation is a 'floor effect' in the test results; that the instrument was not calibrated to pick up changes for the large group of weaker learners.

When we dug a little deeper, we found that among the weakest 40 per cent of learners the Reading Catch-Up Programme intervention may have actually had a small negative effect. Academically stronger learners, however, made real gains in intervention schools.

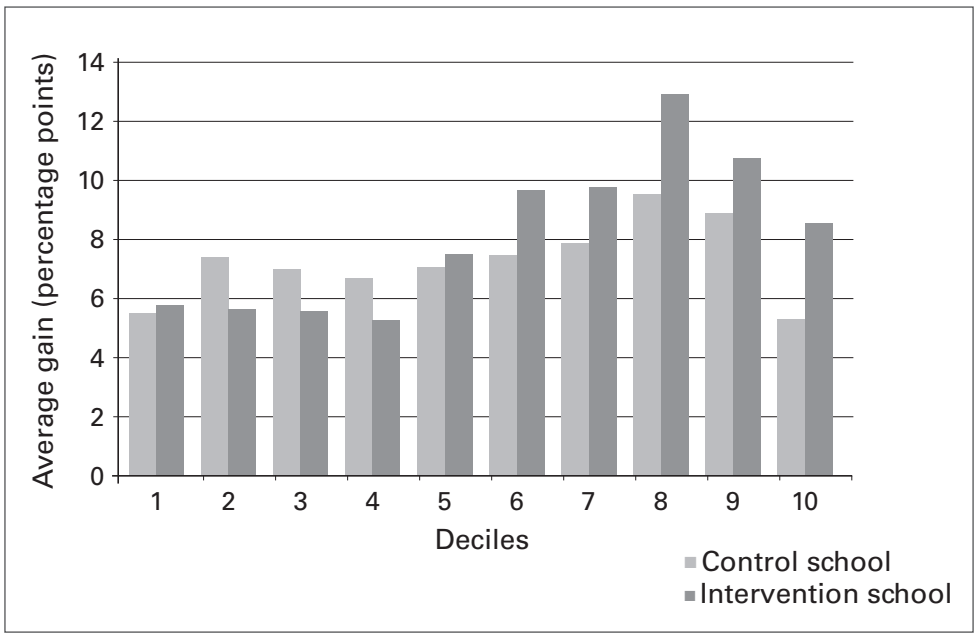


Figure 8.4: Average gain by decile of baseline performance for treatment and control schools
Source: Fleisch et al (2017a)

Unlike the original pre-test/post-test study (reported in Chapter 7), the Pinetown trial found mixed evidence. There was no real difference in the two groups of schools on the endline test. However, the programme benefited academically stronger learners. And a few months on, the children wrote the 2014 Annual National Assessment test. When we analysed the ANA results, we found that schoolchildren in the intervention schools scored higher on average than the children in the control group schools.

Why didn't the programme work as it had in Gauteng? It may have been too short, given that it took some time for teachers to become accustomed to the new instructional approach. In Gauteng, the approach had been in all schools for at least one year or more at the time that the Catch-Up Programme started. It may be that schools need a longer lead time for some teachers to really begin to understand

and enact the new instructional practices. It is probable that the intervention started at the wrong level and that the pacing was too swift. To use Pritchett and Beatty's (2015) concept, it was an overambitious curriculum. There is certainly evidence to support this explanation. My colleagues and I (Fleisch, Pather & Motilal, 2017) recoded and reanalysed the endline spelling subtest. Our analysis showed that most of the Grade 4 schoolchildren were struggling to spell Grade 2-level English words like *bed* and *mat*.

What is the contribution of the Pinetown study to the body of knowledge about the efficacy of structured pedagogical programmes and large-scale change models? Clearly the instructional change model, the triple cocktail, does not work in all situations. Critical factors seem to be: interventions need to take seriously the principle of 'teaching at the right level' (Banerji & Chavan, 2016), and be of long enough duration to take hold.

The Early Grade Reading Study

Building on the experience gained with the GPLMS and the Pinetown studies, specifically the limitation of this research method, and to make a substantial contribution to better understanding interventions that could improve early grade reading across entire systems, the Department of Basic Education initiated the Early Grade Reading Study. With multi-year start-up funding from the International Initiative for Impact Evaluation (3ie), Dr Stephen Taylor, the Director of Research at the national department initiated a randomised control trial of promising approaches to improving early grade reading. Unlike the original Gauteng initiative (which was not designed as a research study), the DBE's Early Grade Reading Study was planned explicitly to provide rigorous evidence to answer key national policy questions. Two models of large-scale intervention were identified for testing: both versions of the education triple cocktail and a third model — a parent literacy support intervention. Each model either had been or was being implemented in South Africa. Given the Gauteng experience with the education triple cocktail and the criticisms associated with the high cost of coaching, the EGRS team recognised the value of testing a 'lite' education triple cocktail, a less expensive version, where coaching was replaced with conventional centralised training.[7]

7 Subsequent to the commencement of the Early Grade Reading Study in 2015 (now referred to as EGRS I), the study team received funding for a second multi-year multi-intervention randomised trial with a focus on reading in English as the second language. While the majority of children in South Africa start learning in their home language, by the time they get to Grade 4, the majority of schools shift to English as the language of instruction. As such, ensuring that all children become fluent readers and writers of English by the end of Grade 3 is an imperative for success in the higher grades. The EGRS II study then was initiated to test the transfer of the basic triple-cocktail model to a second 'subject', ie the second language. The study team also used the opportunity of the new randomised control trial to test the potential effectiveness of using ICT technology with teachers, building on innovations that were taking place in Kenya (Piper et al, 2016).

A number of features made the EGRS randomised control trial distinctive. The study itself was initiated and led by government researchers, but incorporated the Human Sciences Research Council and university partners. Similar studies in the Global South have tended to be driven by institutions in the North. The government/university partnership ensured a clear pathway to policy impact, at the same time guaranteeing dissemination into the wider academic community. Funding for the study came from not one but a group of funders, including UNICEF, Zenex Foundation, the Anglo American Chairman's Fund and the Department of Planning, Monitoring and Evaluation.

The study prioritised both best-practice research design and methodological innovation. The study used a computerised lottery to assign schools to study groups and then tested to ensure that the groups were equivalent. Sample size was statistically determined to ensure that an estimate of the programme impact was within a narrow enough band of uncertainty. Although it was not possible for the study to be 'double blind', separating the intervention implementation from the evaluation meant that the field researchers doing the learner testing did not know if the schools were in the control or intervention groups. In addition to collecting learner outcome data in four separate waves, the study collected classroom instructional practice data from 60 schools, 20 control schools and 20 each of the school-based interventions. The addition of instruction observational data provided insight into the key intervening (or mediating) variables. Finally, using in-depth school case studies, evidence was collected that allowed the study to unpack the actual mechanisms associated with instructional change.

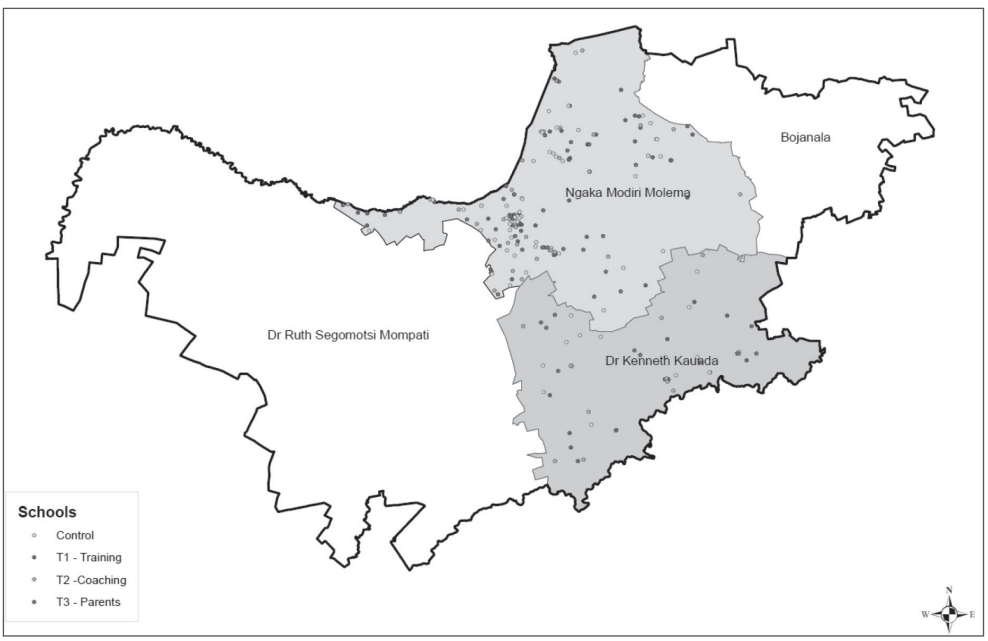


Figure 8.5: Distribution of EGRS schools by group in the North West, 2015–2019
Source: Department of Basic Education (2017)

Intervention implementation and baseline testing began more or less at the same time at the beginning of 2015 in 230 Setswana language schools in two districts in the North West province. By the middle of the year, teachers in the coaching group had received all the learner materials, were using the daily lesson plans and were receiving monthly visits from the coaches. The teachers in the training group were also using the lesson plans, the learner materials and had attended a two-day training session at a local hotel. In the parent literacy intervention, community reading coaches had been recruited from the local communities and had been trained. They had begun to work with parents and started helping them understand how they could support their children's reading development at home. The three intervention school groups continued with their respective programmes into 2016 with the children (or parents) who had been promoted into Grade 2. At the end of 2016, learners were tested individually again by the study's field workers to ascertain their progress in Setswana reading and writing.

The main finding from the study at the end of two years was that the 'coach' education triple-cocktail intervention that included providing lesson plans, quality learning materials and on-site instructional coaching substantially improved school-children's Setswana reading skills. It found that the two other interventions of 'training' and 'parent support' might be working but we could not be completely confident that the improvement was not a chance occurrence. The magnitude of the gains in

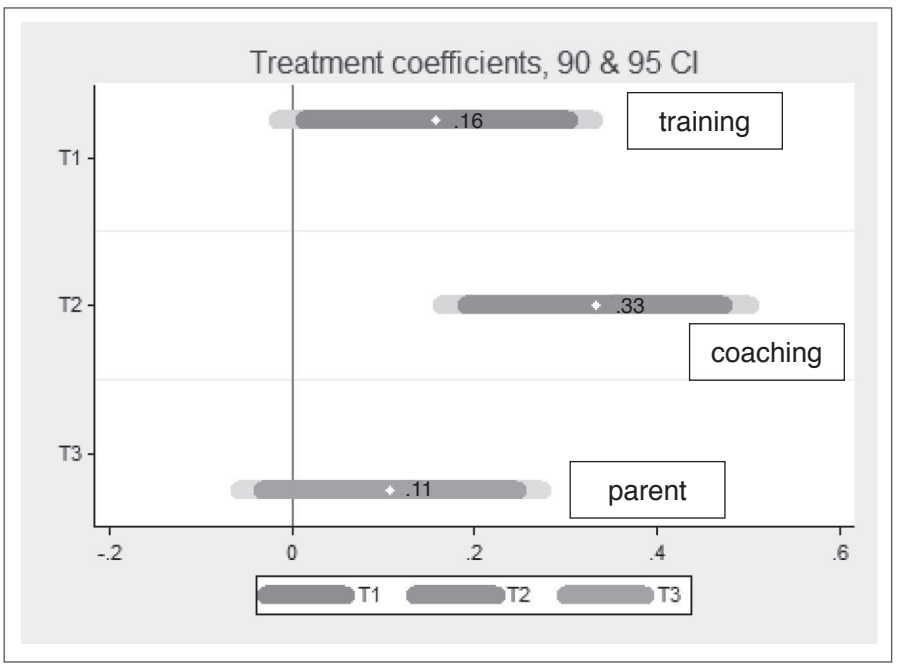


Figure 8.6: EGRS I Year 2 main effects
Note: Grade 1 repeaters excluded
Source: Department of Basic Education (2017)

the 'coach' group was significant. The study estimated that schoolchildren in schools that were randomly selected for the coach intervention were on average 40 per cent of a year of learning ahead of their equivalent counterparts in the control schools.

In Figure 8.6, the solid bars represent statistically significant results at the 90 confidence interval; the bar that crosses the zero intercept indicates that the intervention (parent support) may be effective but that we could not be altogether certain that the reading score gains were not a chance occurrence.

These results are statistically significant, but how educationally meaningful are they? One way to answer this question is to compare the standard deviation effect sizes to equivalent experiments. The International Initiative for Impact Evaluation (3Ie) (Snilstveit, 2016) summary review of structured pedagogic programmes evaluations provides important clues. They found that the average (standardised mean difference) of language programmes that used a structured pedagogic approach was 0.23 of a standard deviation. In the Abdul Latif Jameel Poverty Action Lab (JPAL) study of mathematics interventions, they estimate that mathematics interventions that achieve a change of bigger than 0.4 of a standard deviation should be regarded as 'very large' (*The Economist*, 22 July, 2017).

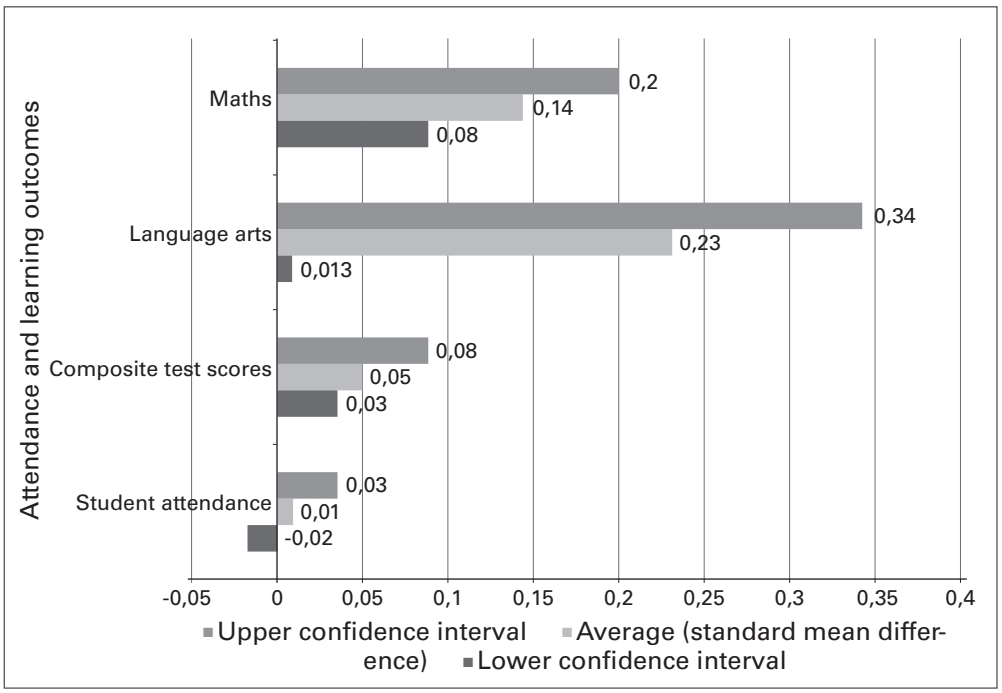


Figure 8.7: Average effect sizes for structured pedagogic programmes across all learning outcomes
Source: Snelsviet, 2016

Table 8.4 shows the actual average scores of each intervention group and the control group on the letter recognition, work recognition, non-word reading and paragraph reading subtests. At the end of 2016, an average Grade 2 learner in the control group could correctly identify 42 letters and 21 words. The equivalent average learners in a school that had received the triple cocktail with coaching could correctly identify 48 letters and almost 26 words. On average, children's paragraph reading in the 'coach' schools was 20 per cent better compared to paragraph reading in the control schools.

Table 8.4: Group subtest average scores for Year 2, 2016

Intervention Group	n–learners	Letter	Words	Non-words	Paragraph
Control	1093	41.9	21.0	15.1	27.5
Training	665	41.7	21.8	16.2	29.2
Coaching	679	47.6	25.7	19.6	34.6
Parents	688	42.6	20.4	15.3	27.2

Notes: Excludes Grade 1 and not specified

While the composite .33 standard deviation suggests that the instructional reform initiative is having a large impact, is the improvement consistent across all reading domains or skills? Figure 8.8 indicates that it is. Schoolchildren in the 'coach' group outperformed the other two intervention groups in letter recognition, word recognition, non-word reading, paragraph reading, comprehension and writing. There also seems to be a cross-over effect on the English subtest items. The only domain or skills that the coach group did less well in was phonological awareness, possibly in part because the parent support programme really focused on phonological awareness skills for parents. The other positive findings are that the interventions do not seem to crowd out other subjects, such as mathematics, as the intervention groups outperformed the control on these test items.

Both relatively strong and relatively weak learners (as measured at baseline) benefited from being in schools that got the 'coach' triple cocktail. Even though the number of schoolchildren who could not correctly identify any letters was roughly the same for the control and coach schools (6–7 per cent of learners), a substantially lower proportion of learners in these 'coach' schools could not correctly identify any non-words and/or read any words in the paragraph test.

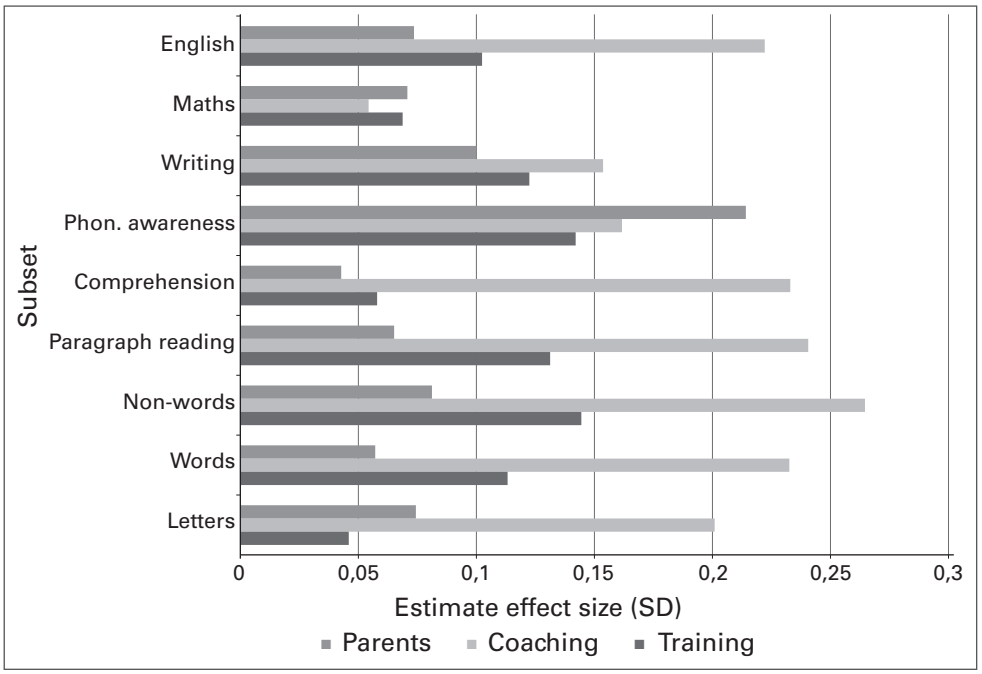


Figure 8.8: EGRS 1 Year 2 Impact on subtests

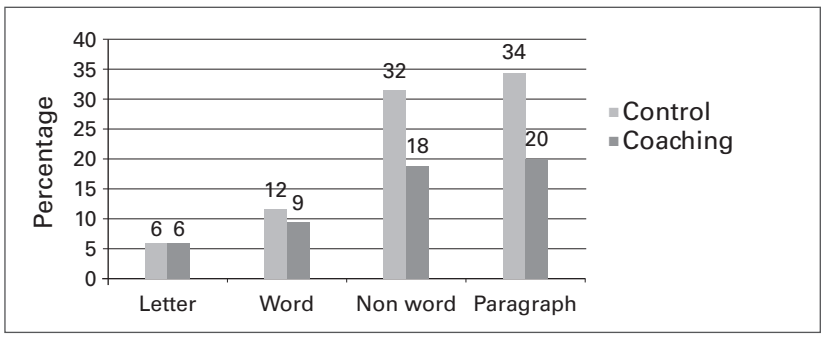


Figure 8.9: Percentage of learners scoring zero on letter, word, non-word and paragraph subtests

When we explored the aggregate school performance on the paragraph reading, we saw that the coach intervention's impact was substantial at the top end of the school sample distribution. The difference between the proportions of schools with average scores between 46 to 51 points is large. Only 5 per cent of the control group school was reading very well (at the top of the subtest range) compared to 18 per cent of schools that had the coach intervention.

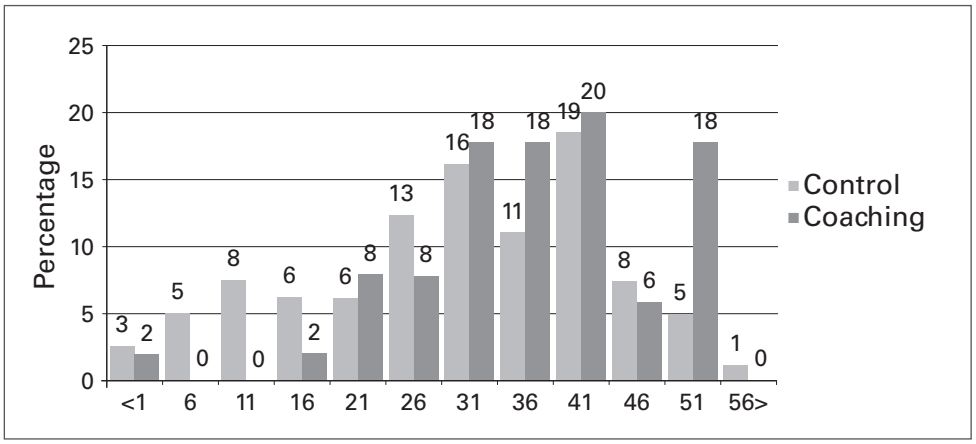


Figure 8.10: Distribution of school average scores for paragraph reading, control and coaching groups

Conclusion: What we know

The research studies on system-wide instructional reform described in this chapter provide a variety of insights. However, for the purpose of this book, I focus on the main question that has animated my research in the last decade: Can large-scale structured pedagogical programmes really work to make change at the instructional core and ultimately raise learning outcomes? The multi-study findings that are presented in this chapter provide compelling evidence that the 'coach' education triple cocktail, does indeed work.

This conclusion is based on rigorous research. The findings are not just from one urban province, but similar findings are now evident both in Gauteng and the North West province. These two contexts are very different: the first is overwhelmingly urban and multilingual while the second is predominantly rural and monolingual. The 'coach' version of the system-wide instructional reform model works not only for the teaching of reading in the Home Language, but evidence has been presented that shows that it also works for early mathematics. And if the 2014 Gauteng ANA results are to be believed, it works at all levels of the primary school. Although the Gauteng natural experiment/RDD and Pinetown studies show that the intervention models worked well for academically stronger learners (possibly because of the lesson plans and reading materials pitched at their level), the EGRS study reveals that school children across the spectrum have and would benefit from being part of a 'coaching' triple-cocktail intervention.

What makes the evidence that much more compelling is that it is not one or two skills that appear to be affected, but as the EGRS study shows, the 'coach' triple-cocktail intervention improves all or nearly all of the key skills associated with effective reading. We also now know with certainty that it is not just the lesson plans

and the learning materials that matter, but that conventional centralised training does not seem to be nearly as effective as in-school instructional coaching. Finally, while the Pinetown studies and the EGRS involved a fair number of schools (groups of between 40 and 50), the Gauteng experience, working in over 1 000 schools, demonstrates the feasibility of rolling out such an intervention across an entire system.

Chapter 9
Conclusion

The starting point of this book is the evidence that the majority of South African schoolchildren are not learning to read and write adequately in primary schools. Whether compared to international middle-income country counterparts or measured against our own national curriculum expectations, our primary schools are not providing children with the necessary learning opportunities that would allow them to flourish.

When we probe below the national averages, the persistent problem of unequal outcomes surfaces with a vengeance. Middle-class children, both black and white, are learning to read and do mathematics, but the vast majority of children growing up in townships, informal settlements and rural areas get to the end of primary school without a solid foundation of core skills and knowledge. Although home and community factors experienced by poor and working-class children often make learning more challenging, a substantial part of the learning failure stems directly from how and what children are taught in our nation's classrooms. Just as we have a bimodal pattern of achievement, so too do we have two distinct systems of classroom instructional practice. In large part, it is the kind of instruction that accounts for the failure of the majority of learners to thrive academically.

The pivot away from social background as the main explanatory variable for learning failure towards a focus on classroom instructional practice opens up real possibilities for change. More government inputs are only partly the answer. While the numerous policies that the South African government has been implementing in the past two decades have certainly made a difference in the lives of millions of children (possibly the best example is the school feeding programme), they have not substantially shifted the unequal patterns of primary school classroom instruction.

What then would it take to shift classroom instructional practices in the vast majority of primary schools in South Africa? This is the problem that has animated this book. Building on the findings of the Gauteng experience and our robust experimental research programme, and paying close attention to understanding the mechanisms of change, this book has presented the case for change at the instructional core and a structured pedagogic programme model of change — what has come to be referred to as the education triple cocktail in South Africa. Converging on interlocking, mutually reinforcing components of daily lesson plans, high-quality

education materials and training/on-site coaching, the book has presented compelling evidence of the efficacy of this approach.

But what will it take to translate this powerful new change knowledge into sustained system-wide reform? Given the complexities and the scope of this problem it would take another full-length book to do it justice. That said, I would like to sketch what I see as the four key challenges. First, we need to understand that the change process is a journey not an event. Second, system-wide change requires us to recognise that although substantial advances have been made on lesson plans, materials and coaching, much more needs to be learned to make the cocktail more efficacious. Third, while the focus is on the instructional core, we cannot ignore problematic structural constraints that can only be addressed by government policy intervention. Finally, to make change at this scale, with this level of complexity, requires real system-level leadership.

Change is not an event, it is a journey

There is a strong temptation to treat change as linear and technical. We could diagnose the problem, develop a prototype intervention, pilot it in a few schools and iron out the kinks, test it in a large representative sample in a randomised control trial and if shown to be cost-effective, roll it out at scale. Problem solved.

It does not work like this because the change problem demands a multi-stage, multi-phase and multiple pathway approach. And each stage, phase and pathway has its own unique dynamics. The idea of change as a journey, although a cliché, remains a powerful way of understanding these dynamics.

Scaling-up

The first obvious change problem is moving from the large representative sample experimental trials that may have involved between 100 and 300 schools, to planning for system change in thousands of schools. It is often not just a case of an exponential increase in the number of schools, teachers and learners, but would involve moving from trialling in one subject in one grade at a time in a few districts, to implementation in multiple subjects, for example, Home Language and First Additional Language, multiple languages (in Gauteng this would mean eight African languages), in more than one province and in multiple grades, that is, Grades 1 to 3 concurrently. Scaling up also forces plans to confront the challenges of constrained markets. In the randomised control trial it was relatively simple to recruit a handful of qualified coaches, but when working at scale, recruitment of qualified and experienced staff is bound to be difficult.

Beyond the plateau

The introduction of an education triple cocktail or any structured pedagogic programme across a poorly performing education system or sub-system is best

seen as an intervention for the first or initial stage of a multi-stage change journey. Improvements teachers can see in their classrooms restore hope and belief in the possibility of change. But without careful, systematic planning and implementation of the next stage in the improvement journey, early gains can plateau and even reverse. Subsequent stages of planned improvement require institutionalisation and a deepening of new instructional practices. For many teachers, simply enacting the planned lesson would not be enough after a few years. This is where enhancing teachers' knowledge becomes critical. Once on the road to improvement, many teachers will need to acquire a deeper understanding of theory and research on how children learn to read and why certain techniques and methodologies work. Acquiring technical expertise, sharing good practices in a local community of practitioners alongside a renewed sense of hope, would allow teachers to extend themselves beyond the basic lesson plans that have been designed as a platform on which continuous improvement would be built. District subject advisers and school-based leaders, both heads of departments and principals, need to be activated and capacitated during this institutionalisation stage of change. Not only do subject advisers and school managers need to acquire a deep understanding of the new instructional practices, but they need to ensure that the enabling conditions for institutionalisation are in place. These enabling conditions include the maintenance and expansion of appropriate and effective educational resources and the recruitment and induction of new teachers into the new instructional practice. While performance accountability is *not* a key feature of the education triple cocktail in the initial stage of change, as management by performance targets requires a high level of capacity and trust, both of which are in short supply in the current system, there is certainly a place for it during institutionalisation.

During the initial stage, districts and head office professionals need to ready themselves to take on new roles with regard to the new instructional practice. While external coaches have the advantage of being able to focus exclusively on the new instructional core and have the potential to build trust, one-on-one coaching is probably not a sustainable and cost-effective option in the medium and long term. This means we need to build and test models of district and within-school support that focus on continuous improvement.

Effective change models do not work in all schools

The intervention model, the education triple-cocktail consists of a strong, effective form of professional development — instructional coaching, tightly coupled with direct guidance to teachers on the daily business processes of teaching, and the right learning resources and materials. However, as simple and powerful as these components are in combination, it does not and will not work to shift instruction in *all* township and rural schools. In the North West province's Early Grade Reading Study, the results showed very large gains for the education triple cocktail

in more urban primary schools but the successful intervention model was less effective in deep rural settings (Taylor, 2017). Taylor found that factors which limited impact in rural settings included long travel distances, which prevented attendance at programme activities and coaches' access to schools, and lost teaching time (possibly caused by weaker accountability of schools to government and parents).

For genuinely dysfunctional primary schools, characterised by poor infrastructure, weak leadership, high teacher and learner absenteeism and low morale, a structured pedagogic intervention is unlikely to lead to sustained changes in instructional practice. In these schools, teachers often simply do not do the programme. Even with coaches and lesson plans, the educational materials remain in the classrooms unopened. If teachers are 'tired' (marking time until they retire) or unwilling to exert effort, different kinds of interventions need to be developed to create the right pre-conditions for instructional change.

Move into Intermediate Phase

Because of the cumulative effect of limited learning, the education triple cocktail is unlikely to have much impact if implemented in higher grades concurrently. If schoolchildren's reading and mathematics skills are two or more grades below the curriculum expectations at the end of the Foundation Phase, even the most skilled teachers using the most effective new instructional programmes are unlikely to remedy system-wide learning backlogs at the same time as covering the grade-specific curriculum.

Change at the instructional core requires us to plan for the multi-stages of change and plan different approaches depending on the change-readiness of schools. We also need to plan for improvement at the intermediate and senior school phases. If the education system intends raising the bar and narrowing the learning gap, then we need to ensure these early grade learning gains become permanent. To do this we will need to plan to improve teaching in the middle-school years and beyond. The teaching practices enacted by many Intermediate and Senior Phase teachers in township and rural schools have evolved to take into account the fact that schoolchildren struggle to learn to read. The dominant oral teaching approaches would need to shift to a greater emphasis on using books and other reading resources for learning. Intermediate Phase teachers would need to continue the teaching of reading and writing, albeit at more advanced levels of interpretative and critical literacy skills.

R&D and the triple cocktail components

The education triple cocktail with the three interconnected and mutually reinforcing components has shown to be a powerful stimulus in the initial stage of the instructional change journey for many township and rural schools. Given the pace

at which implementation has occurred, insufficient time and resources were committed to research and development. More R&D is needed to improve each of the components and how they work as an aligned and seamless instructional infrastructure.

Lesson plans

Notwithstanding lingering doubts, there is a growing consensus in the research and practitioner communities about the usefulness of daily lesson plans for teachers in transforming their instructional practices. That said, there is a real need to understand how teachers use lesson plans to change their practice and how the plans can be improved to make the work of teaching easier and more effective.

Ben Piper and a team at the Research Triangle Institute recently (2018) did a comparative study of teacher guides and lesson plans in 19 projects and 13 African countries. Building on their findings, they made a number of useful recommendations. For example, for lesson plans to be used effectively, they need to be written in the language of instruction. There should be heavier scripting at the beginning of the year and once routines have been established, reduced to lighter scripting. Good lesson plans should limit the number of activities per lesson for the day, provide the pacing or time for each activity, incorporate embedded checks for understanding in each lesson and provide guidance for differentiated instruction. From the perspective of teachers, lesson plan guide books need to be light weight and easy to use. That said, they should contain suggestions of ways in which teachers can extend the lesson activities and embed the 'why' of the activities in the lesson plans (Shalem et al, 2016).

There is equally a need for more 'out of the box' research on lesson plans that go beyond the old conventional paper and print variety. Currently, in the Second Early Grade Reading Study (Kotze et al, 2018) under way in Mpumalanga, the research team is testing the practicability and effectiveness of putting lesson plans on electronic tablets. The electronic format would allow teachers to access linked electronic media resources that model the new practices and provide other opportunities for teachers to learn the 'why' behind the lesson activities.

Beyond the form of the lesson plans, there is certainly demand for research on the content of the core methodologies embedded in the activities. For example, we need to know more about best-practice teaching of phonological awareness, phonics and reading fluency in all South African languages. Current research on African language structures for early grade reading (Spaull, Pretorius & Mohohlwane, 2017) has highlighted the need to revisit the design of language activities and more closely align them with the linguistic structures and form of each of the languages or language groups. While this study was designed to contribute to establishing oral reading fluency norms, it highlighted the huge difference in language structures between languages, differences that signal the need to reconsider the appropriateness

of phonological awareness approaches that have been largely translated straight from English.

Educational materials

Educational materials clearly matter. The quality and quantity of reading materials that children have access to in the early grades are particularly critical. The CAPS curriculum suggests the need for four kinds of early grade reading materials: read-aloud texts, shared-reading texts to model the reading process (eg big books), group-guided, or levelled readers, and books for independent reading. For group-guided reading, the typical resource packages in middle-class schools were commercial reading schemes such as Oxford Reading Tree or Ginn 360. With the exception of the Vula Bula books (Katz, 2013), until recently, few equivalent resources of quality were available in African languages. While it is now widely recognised that back-translation from English to African languages is not appropriate, questions remain about whether collections that include only 6 or 8 copies per title, with 20 or 30 titles for each grade, are sufficient in large classes to promote reading for meaning. Research is also needed on the optimal balance of decodable texts and predictable texts for early grade reading in African languages (Sailor et al, 2009). We also need to understand how to provide book collections for independent reading, the kind of reading that can move children from merely acquiring the skills of reading to help them internalise the habit of reading.

Coaching

Instructional coaching is now understood as a powerful way to help teachers to start enacting new practices (Kraft et al, 2016; Popova et al, 2016) and this enactment is the necessary first step to changing beliefs and attitudes about instruction (Guskey, 2002). In the EGRS I study (Cilliers et al, 2018), we found that traditional face-to-face training in a centralised venue is less effective than in-school coaching. Albornoz et al (2017) in Argentina found substantial learning gains when learners were in classes with teachers, particularly for inexperienced teachers being trained using structured programmes and coaching. These findings build on the Piper and Zuilkowski (2015b) coaching study that provides valuable insights on the relative cost-effectiveness of different coach-teacher ratios in Kenya.

But too little is known about the internal workings of the coaching process itself. We also need to know if there are alternatives to the relatively high cost of hiring external personnel to do the coaching, and if district subject advisers or heads of department could coach if their work was restructured and they were properly coached and trained themselves. We also need to do more research on how we can incorporate ICT into the coaching process, as in Kenya's Tusome initiative and the Second Early Grade Reading Study.

While the prime purpose of coaching is to enable site-based learning, there is

a second and equally important role that coaching plays in instructional change. Since the Hargreaves (2005) and Fullan (2007) studies on educational change, there has been a growing recognition of the significant impact, both positive and negative, that emotions have on educational change. We need to know more about the role that coaches can play in helping to manage the emotional labour of change.

Policy

As powerful as the education triple cocktail might be in kick-starting instructional change at scale, there are certain binding constraints (things that must be tackled first, as they preclude progress) (Van der Berg et al, 2016). The Stellenbosch based group (ReSEP), have suggested four such constraints: weak institutional capacity, undue union influence, weak teacher knowledge and wasted time. While these are all undoubtedly important, I would point to two policy gaps that impede system-wide instructional change.

Very large classes

Research in the Education and Social Economic Policy study (Spaull et al, 2016) on excessive early grade class sizes in no-fee primary schools found that, in the Eastern Cape and Limpopo, between 10 per cent and 15 per cent of Grade 1–3 learners were in classes of more than 60 learners, and in five of the nine provinces more than 30 per cent of Grade 1–3 learners attended classes of more than 45 learners. Although the learner-teacher ratios have been guided by the Education Labour Relations Agreement signed in 1995 at a ratio of 40:1 in primary schools, the actual primary school class sizes have been consistently larger. We need to know more about what drives large and very large class sizes and how policy can help. Irrespective of the causes, there can be little doubt that system-wide instructional reform would be all but impossible in the significant portion of the system with excessively large classes. While the infrastructure backlog is clearly not as widespread, when school buildings, furniture and supplies do not comply with the minimum norms and standards, their absence represents a real binding constraint.

Universal assessment

As system-wide instructional change gains momentum, standardised external assessment of learning takes on a vital role. All stakeholders in the change journey need to know if learning system-wide is getting better. The best way to do this is to show that across the education system more schoolchildren are mastering the basic skills and are able to read for meaning. As Dan Wagner (2003) pointed out, governments need system-wide learning assessments that are smaller, quicker and cheaper. We need to find the balance between the simplicity of the Annual Status of Education Report (ASER) model (Banerji & Chavan, 2016), which tests rural children's letter

and word recognition and sentence and paragraph reading, and the more complex but diagnostically useful Early Grade Reading Assessment (EGRA). We also need to make certain that the assessment or test instruments are valid, reliable and that the results are comparable over time. But without this information, we have neither the information to judge whether the improved learning is real, nor, more importantly, information to help do data-driven decision-making.

At a policy level, we need to be mindful of the risks of standardised assessments. At the earliest stage of the change journey with the education triple cocktail, too much data could have had perverse effects. The unhealthy preoccupation of the education system with comparative performance is evident from the Grade 12 National Senior Certificate examination and even the ANA (Kanjee & Moloi, 2014). Everyone wanted to show how they had improved on the standardised matrices. Gaming standardised assessment is hard to resist at all levels, from the teachers, schools, districts and provincial departments to the national government.

Leadership

Many scholars writing about system-wide instructional reform have pointed to the centrality of leadership for successful change. Levin and Fullan (2008) talked about inspiring leadership and building a guiding coalition from the experience of provincial educational reform in England and Canada. Mourshed and colleagues (2010) from McKinsey point to lessons from a diverse set of country systems and how new political or strategic leaders have more often than not been the drivers of change across their systems. Most recently the World Bank (2018), in their report, *Learning to Realise Education's Promise,* highlights the role of leadership in moving systems past misalignment and incoherence.

Political leadership

As a tool of system-wide instructional improvement, the structured pedagogic programmes or education triple cocktail specifically assume that moving forward on change requires decisive political leadership and the substantial resources and authority of government. This approach is the opposite of bottom-up change, which relies on informal social networks of teachers sharing ideas. The advantage of an informal approach to change is that the teacher-agents are highly motivated and often are evangelical about the new approach. The disadvantage is that new practice tends to remain concentrated in the hands of a few enthusiasts.

This structured pedagogic model of change is geared neither to the enthusiasts or the laggards but to the solid majority of teachers who will follow a vision if it is well designed, makes sense and is part and parcel of a government policy and programme. In this context, political leadership in the change journey plays a number of critical roles. Leaders must articulate the vision and remain steadfastly

on-message in communicating that vision to all education stakeholders, particularly teachers, parents, department officials and the press.

The potential danger with political leaders driving the change is well documented and is associated with political or policy 'churn', that is, each successive new political leader needs to be the innovator or originator, and wishes to have his or her 'legacy' attached to a successful initiative. As change of this magnitude will take at least a decade in systems such as those of South Africa, political churn and legacy projects represent a real threat.

There is another form of system-leadership that is critical, namely, the leadership of and by top and senior public service managers. While political leaders can provide the vision and play a vital role in communicating the message, it is the directors in provincial and national departments who do the detailed planning that is needed to put into place complex multi-component interventions such as the education triple cocktail. The danger within the public service is that the initiatives are too often seen an 'add-ons' or vanity projects. Senior managers adopt an attitude 'that this too will pass in time'. Unless top and senior managers buy in and do the hard work of aligning budgets, programmes and initiatives with the day-to-day work streams of the organisation, complex changes around the instructional core in classrooms may start but will not be sustained or institutionalised. Each branch and each directorate associated directly or even indirectly with the instructional core needs to restructure to support the new instructional core. In particular, directorates responsible for the curriculum, assessment, strategic planning, finance, teacher development and educational materials need to substantially revise their business processes and ready themselves for alignment.

Beyond government, scaling up the education triple cocktail requires alignment with and the support of a range of stakeholders in the wider society. While a great deal of emphasis has been placed on the potential difficulties associated with partnerships with teacher unions, because of the existence of structured spaces such as the national and provincial education labour relations councils, these can be managed. Less easy to manage are the non-governmental agencies and the funders or donors. The NGOs often have their own distinct 'model' and wish to advance their own 'brand'. At school level, teachers need to navigate their way through a myriad of NGO programmes, often speaking at cross purposes. There is a similar problem with external funders and donors. Priorities set in corporate boardrooms need to be aligned to the overall objective of system-wide instructional change. This is why Levin and Fullan talk about the need to build 'guiding coalitions', a grouping of key decision-makers who come together to maximise coherence and minimise distractions.

Bibliography

Albornoz, F., Anauati, M. V., Furman, M., Luzuriaga, M., Podestá, M. E. & Taylor, I. (2017). *Training to teach science: experimental evidence from Argentina*. Centre for Research in Economic Development and International Trade, University of Nottingham. https://www.nottingham.ac.uk/credit/documents/papers/2017/17-08.pdf. (Accessed 14 May 2018).

Ball, D. L. & Cohen, D. K. (1996). Reform by the book: What is: or might be: the role of curriculum materials in teacher learning and instructional reform?. *Educational Researcher*, *25*(9): 6–14.

Ball, D. L. & Cohen, D. K. (1999). Instruction, capacity, and improvement. Philadelphia, PA: Consortium for Policy Research in Education, University of Pennsylvania.

Ball, D. L. & Forzani, F. M. (2007). Wallace Foundation Distinguished Lecture—What Makes Education Research 'Educational'?. *Educational Researcher*, *36*(9): 529–540.

Banerjee, A. V., Banerji, R., Duflo, E., Glennerster, R. & Khemani, S. (2010). Pitfalls of participatory programs: Evidence from a randomized evaluation in education in India. *American Economic Journal: Economic Policy*, *2*(1): 1–30. Retrieved from www.aeaweb.org/articles.php?doi=10.1257/pol.2.1.1. (Accessed 14 May 2018).

Banerjee, A., Banerji, R., Duflo, E. & Walton, M. (2012). Effective pedagogies and a resistant education system: Experimental evidence on interventions to improve basic skills in rural India (Unpublished manuscript).

Banerjee, A. & Duflo, E. (2012). *Poor economics: A radical rethinking of the way to fight global poverty.. New York:* Public Affairs.

Banerjee, A., Banerji, R., Berry, R., Duflo, E., Mukherji, S. & Walton, M. (2015). Teaching at the right level: Evidence from randomized evaluations in India. NBER Working Paper No. 22746. Cambridge, MA: The National Bureau of Economic Research.

Banerji, R. (2014). An intervention improves student reading. *Phi Delta Kappan*, *95*(6): 74.

Banerji, R. & Duflo, E. (2015). *'Teaching at the right level': Solutions for low learning levels in India*. (Blog). Retrieved from www.ideasforindia.in/article.aspx?article_id=1541#sthash.PXKp4KAe.dpuf. (Accessed 14 May 2018).

Banerji, R. & Chavan, M. (2016). Improving literacy and math instruction at scale in India's primary schools: The case of Pratham's read India program. *Journal of Educational Change*, *7*(4): 453–475.

Bapna, A. & Sharma, N. (2015). School education in India: A handbook. *EvalDesign*. Retrieved from www.evaldesign.com/uploads/2/3/8/2/23823775/school_education_india-evaldesign.pdf. (Accessed 14 May 2018).

Barber, M. (2008). *Instruction to deliver*. London: Methuen.

Beatty, B. (2011). The dilemma of scripted instruction: Comparing teacher autonomy, fidelity, and resistance in the Froebelian Kindergarten, Montessori, Direct Instruction, and Success for All. *Teachers College Record*, *113*(3): 395–430.

Berman, P. & McLaughlin, M. W. (1974). Federal Programs Supporting Educational Change: A Model of Educational Change. Volume I. https://files.eric.ed.gov/fulltext/ED099957.pdf. (Accessed 14 May 2018).

Besharati, N. A. (2016). Evaluating development effectiveness: assessing and comparing the impact of education interventions in South Africa. (Doctoral dissertation, University of the Witwatersrand).

Biputh, B. & McKenna, S. (2010). Tensions in the quality assurance processes in post-apartheid South African schools. *Compare*, *40*(3): 279–291.

Bizos, E. N. (2010). Authoring lives: a case study of how grade 6 children in a South African township school construct themselves as readers and writers. (Doctoral dissertation, University of the Witwatersrand).

Botha, D. (2007). Early Childhood Literacy and Emergent Bilingualism: Vocabulary Teaching and Learning in One Soweto Classroom. (Honours thesis, University of the Witwatersrand).

Bunyi, G., Cherotich, I. & Piper, B. (2012). *Primary math and reading (PRIMR) program: Kenya.* Research Triangle Park, NC: RTI International. Retrieved from http://pdf.usaid.gov/pdf_docs/PA00HX76.pdf. (Accessed 14 May 2018).

Carnoy, M., Chisholm, L. & Chilisa, B. (eds). (2012). *The low achievement trap: Comparing schooling in Botswana and South Africa*. HSRC Press. Cape Town.

Carnoy, M., Gove, A. K. & Marshall, J. H. (2007). *Cuba's academic advantage: Why students in Cuba do better in school*. Standard, CA: Stanford University Press.

Case, A. & Deaton, A. (1999). School inputs and educational outcomes in South Africa. *The Quarterly Journal of Economics*, *114*(3): 1047–1084.

Cilliers, J., Fleisch, B., Prinsloo, C., Reddy, V. & Taylor, S. (2018). *How to improve teaching practice? Experimental comparison of centralized training and in-classroom coaching.* http://media.withtank.com/45bb06169c/egrs.pdf. (Accessed 6 March 2018).

Cilliers, J., Fleisch, B., Prinsloo, C. & Taylor, S. (2016). Scripted lesson plans and teaching to the level of the child: A randomized evaluation of two interventions aimed at improving early-grade reading in South Africa. (Working Paper).

City, E. A., Elmore, R. F., Fiarman, S. E. & Teitel, L. (2009). *Instructional rounds in education: A network approach to improving teaching and learning*. Cambridge, MA: Harvard Education Press.

Cohen, D. K. (2011). *Teaching and its predicaments*. Cambridge, MA: Harvard University Press.

Cohen, D. K., Raudenbush, S. W. & Ball, D. L. (2003). Resources, instruction, and research. *Educational Evaluation and Policy Analysis*, *25*(2): 119–142.

Cohen, D. K. & Spillane, J. (1993). Policy and practice: The relations between governance and instruction. In S. Fuhrman (ed.) *Designing coherent education policy: Improving the system*. San Francisco: Jossey-Bass.

Condy, J. & Donald, D. (2005). The impact of two years of CLE literacy interventions. In K. Parry (ed) *Teaching Reading in African Schools*. Kampala: Fountain Publishers.

Conn, K. M. (2017). Identifying Effective Education Interventions in Sub-Saharan Africa: A Meta-Analysis of Impact Evaluations. *Review of Educational Research*, *87*(5): 863–898.

Cremin LA. (1976) *Public education*. New York: Basic Books.

Crouch, L. (2005). South Africa equity and quality reforms: possible lessons. *Journal of Education for International Development*, *1*(1): 1–18.

Crouch, L. (2011). Motivating early grade instruction and learning: Institutional issues. In Gove, A. & Wetterberg, A. (eds) *The Early Grade Reading Assessment: Applications and Interventions to Improve Basic Literacy*. Research Triangle Park, RTI International.

Crouch, L. & Mabogoane, T. (1998). When the residuals matter more than the coefficients: An educational perspective. *Journal for Studies in Economic and Econometrics*, *22*: 1–14.

Crouch, L. & Mabogoane, T. (2001). No magic bullets, just tracer bullets: The role of learning resources, social advantage, and education management in improving the performance of South African schools. *Social Dynamics*, *27*(1): 60–78.

D'Amico, L. & Stein, M. (2002). Inquiry at the crossroads of policy and learning: A study of a district-wide literacy initiative. *The Teachers College Record*, *104*(7): 1313–1344.

De Clercq, F. (2010). Policy mediation and leadership: Insights from provincial implementers of South African school evaluation policies. *Southern African Review of Education with Education with Production*, *16*(2): 100–116.

Deaton, A. (2010). Instruments, randomization, and learning about development. *Journal of Economic Literature*, *48*(2): 424–455.

Deaton, A. & Cartwright, N. (2016). *Understanding and misunderstanding randomized controlled trials* (No. w22595). Cambridge, MA:National Bureau of Economic Research.

Department of Basic Education (2009). Sepedi 10–20. Systematic Method for Reading Success. Pretoria.

Department of Basic Education (2011). Annual National Assessment 2011: Report on the ANA of 2011. Pretoria.

Department of Basic Education (2014). Annual National Assessment 2014: Report on the ANA of 2014. Pretoria.

Department of Basic Education (2017) Summary Report: Impact of Year 2 Evaluation. Early Grade Reading Study. Pretoria.

Dixon, K. (2010). *Literacy, Power and the Schooled Body: Learning in Time and Space.* New York: Routledge.

Elley, W. & Cutting, B. (2001). The 'Sunshine in South Africa' literacy project. *International Journal of Educational Research, 35*: 193–203.

Elmore, R. F. & Burney, D. (1997). *Investing in teacher learning: Staff development and instructional improvement in community school district # 2.* New York City: National Commission on Teaching & America's Future, Teachers College, Columbia University.

Evans, D. K. & Popova, A. (2016). What really works to improve learning in developing countries? An analysis of divergent findings in systematic reviews. *The World Bank Research Observer*, *31*(2): 242–270.

Fiske, E. B. & Ladd, H. F. (2004). *Elusive equity: Education reform in post-apartheid South Africa*. Washington DC: Brookings Institution Press.

Flanagan, W. (1992). Reading programmes: are they the best form of interventions in our schools?. *Perspectives in Education*, *14*(1): 35–45.

Flanagan, W. (1997). Pedagogy and quality primary education in South Africa. *Education and Development: Tradition And Innovation.* L. J. London: Cassell.

Flanagan, W. (2002). Pedagogical discourse: teacher education programmes and social transformation in South Africa. *International Journal of Educational Development*, *12*(1): 27–35.

Fleisch, B. (2008). *Primary education in crisis: Why South African schoolchildren underachieve in reading and mathematics*. Cape Town: Juta and Company Ltd.

Fleisch, B. (2011). Mona Mourshed, Chinezi Chijioke and Michael Barber: How the world's most improved school systems keep getting better. *Journal of Educational Change, 12*: 469–472.

Fleisch, B. (2016). System-wide improvement at the instructional core: Changing reading teaching in South Africa. *Journal of Educational Change*, *17*(4): 437–451.

Fleisch, B. (2017). Experimental Research and Educational Change: Methodological Insights from the Global South. In H. Malone (ed) *Future of Educational Change.* New York: Routledge.

Fleisch, B., Pather, K. & Motilal, G. (2017). The patterns and prevalence of monosyllabic three-letter-word spelling errors made by South African English First Additional Language learners. *South African Journal of Childhood Education, 7*(1): 1–10.

Fleisch, B. & Schöer, V. (2014). Large-scale instructional reform in the Global South: insights from the mid-point evaluation of the Gauteng Primary Language and Mathematics Strategy. *South African Journal of Education*, *34*(3): 01–12.

Fleisch, B., Schöer, V., Roberts, G. & Thornton, A. (2016) System-wide improvement in early-grade mathematics: new evidence from the Gauteng Primary Literacy and Mathematics Strategy, *International Journal of Educational Development, 49*:157–174.

Fleisch, B., Taylor, N., Herholdt, R. & Sapire, I. (2011) Evaluation of *Back to Basics* Mathematics Workbooks: A Randomised Control Trial of Primary Mathematics Research Project. *South African Journal of Education*, *31*: 488–504.

Fleisch, B., Taylor, S., Schöer, V. & Mabogoane, T. (2017a). Failing to catch up in reading in the middle years: The findings of the impact evaluation of the Reading Catch-Up Programme in South Africa. *International Journal of Educational Development*, *53*: 36–47.

Fleisch, B., Taylor, S., Schöer, V. & Mabogoane, T. (2017b). The value of large-scale randomised control trials in system-wide improvement: The case of the Reading Catch-Up Programme. *South African Journal of Education*, *37*(1): 1–13.

Fuhrman, S. H. (1993). *Designing Coherent Education Policy: Improving the System*. San Francisco, CA: Jossey-Bass.

Fullan, M. (2000). The return of large-scale reform. *Journal of Educational Change*, *1*(1): 5–27.

Fullan, M. (2007). *The new meaning of educational change*. New York: Routledge.

Fullan, M. (2011). *The six secrets of change: What the best leaders do to help their organizations survive and thrive.* Hoboken, New Jersey: John Wiley & Sons.

Fullan, M. (2015). *The new meaning of educational change* (4th ed). New York, NY: Teachers College Press, Columbia University.

Fullan, M. & Earl, L. (2002). United Kingdom national literacy and numeracy strategies—Large scale reform. *Journal of Educational Change*, *3*(1): 1–5.

Fullan, M. & Knight, J. (2011). Coaches as system leaders. *Educational Leadership,* October 2011: 50–56.

Fuller, B. (1987). What school factors raise achievement in the Third World? *Review of educational research*, *57*(3): 255–292.

Fuller, B., Snyder, C. W., Chapman, D. & Hua, H. (1994). Explaining variation in teaching practices? Effects of state policy, teacher background and curricula in Southern Africa? *Teaching & Teacher Education, 10*(2): 141–156.

Gains, P. (2010). Learning about Literacy: Teachers' Conceptualisations and Enactments of Early Literacy Pedagogy in South African Grade One Classrooms. (PhD thesis, University of the Witwatersrand).

Gains, P. & Mfulathela, G. (2005). Breakthrough to Literacy (BTL): solutions for teaching and learning initial literacy in African languages. *Teaching Reading in African Schools.* K. Perry. Kampala: Fountain Publishers.

Gallucci, C., Van Lare, M. D., Yoon, I. H. & Boatright, B. (2010). Instructional coaching: Building theory about the role and organizational support for professional learning. *American Educational Research Journal*, *47*(4): 919–963.

Ganasi, R. (2010). The Reading Experiences of Grade Four Children. (M Ed Research Report, University of KwaZulu-Natal).

Gauteng Department of Education (2012). Intersen Catch-up Programme Lesson Plan Grade 4. Johannesburg.

Glewwe, P. & Muralidharan, K. (2015, October). *Improving school education outcomes in developing countries: Evidence, knowledge gaps and policy implications*. RISE working paper. Retrieved from www.bsg.ox.ac.uk/sites/www.bsg.ox.ac.uk/files/documents/RISE_WP-001_Glewwe_Muralidharan.pdf. (Accessed 26 June 2018).

Guskey, T. R. (2002). Professional development and teacher change. *Teachers and Teaching*, *8*(3): 381–391.

Gustafsson, M. & Patel, F. (2006). Undoing the apartheid legacy: Pro-poor spending shifts in the South African public school system. *Perspectives in education*, *24*(2): 65–77.
Haddaway, N. R., Land, M. & Macura, B. (2016). A little learning is a dangerous thing: a call for better understanding of the term 'systematic review'. *Environment International*, *99*: 356–360
Hargreaves, A. (2003). *Teaching in the knowledge society: Education in the age of insecurity*. New York: Teachers College Press.
Hargreaves, A. (2005). The emotions of teaching and educational change. In *Extending educational change* (pp. 278–295). Dordrecht: Springer.
Hargreaves, A. & Fullan, M. (2012). *Professional capital: Transforming teaching in every school*. New York: Teachers College Press, Columbia University.
Hargreaves, A. & Shirley, D. L. (eds). (2009). *The fourth way: The inspiring future for educational change*. Thousand Oaks, CA: Corwin Press.
Harris, A. & Jones, M. S. (eds). (2015). *Leading futures: Global perspectives on educational leadership*. Thousand Oaks, CA: Sage.
Hatch, T. (2013). Innovation at the Core; R&D. *Phi Delta Kappan*, *95*(3): 34–38.
Hellman, L. (2012). GPLMS Intersen Catch-up Programme: Analysis of Results. Johannesburg.
Herbst, P. & Chazan, D. (2012). On the instructional triangle and sources of justification for actions in mathematics teaching. *ZDM*, *44*(5): 601–612.
Hiebert, J. & Morris, A. (2012). Teaching, rather than teachers, as a path toward improving classroom instruction. *Journal of Teacher Education*, *63*(2): 92–102.
Hiebert, J. & Stigler, J. W. (2017). Teaching Versus Teachers as a Lever for Change: Comparing a Japanese and a US Perspective on Improving Instruction. *Educational Researcher*, *46*(4): 169–176.
Hoadley, U. (2012). What do we know about teaching and learning in South African primary schools?. *Education as Change*, *16*(2): 187–202.
Hoadley, U. (2017). 'Learning to fly': pedagogy in the Foundation Phase in the context of the CAPS reform. *Journal of Education,* 67: 13–38.
Hopkins, D., Harris, A. & Jackson, D. (1997). Understanding the school's capacity for development: growth states and strategies. *School Leadership & Management*, *17*(3): 401–412.
Howie, S., Combrinck, C., Roux, K., Tshele, M., Mokoena, G., & Palane, N. M. (2017). PIRLS 2016. Centre for Evaluation and Assessment, University of Pretoria.
Howie, S., Van Staden, S., Tshele, M., Dowse, C. & Zimmerman, L. (2012). PIRLS 2011: South African children's reading literacy achievement summary report. Centre for Evaluation and Assessment, University of Pretoria.
Howie, S., Venter, E., Van Staden, S., Zimmerman, L., Long, C., Du Toit, C., Scherman, V. & Archer, E. (2008). Progress in International Reading Literacy Study 2006. Centre for Evaluation and Assessment, University of Pretoria.
Hugo, C. (no date). The Business Trust: Learning for Living Project.
Hunt, G. (2006). Failure to thrive? The community literacy strand of the Additive Bilingual Project at the Eastern Cape Community School, South Africa. *Journal of Research in Reading*, *30*(1): 80–96.
Janks, H. (2014). Globalisation, diversity, and education: A South African perspective. *The Educational Forum*, 78: 8–25.
Jansen, J. D. (2002). Political symbolism as policy craft: Explaining non-reform in South African education after apartheid. *Journal of Education Policy*, *17*(2): 199–215.
Kanjee, A. & Moloi, Q. (2014). South African teachers' use of national assessment data. *South African Journal of Childhood Education*, *4*(2): 90–113.
Katz, J. (2013). Vula Bula. *The Development of South Africa's First Graded Basal Readers in Eight African Languages*. South African Education Research Association Meeting. Durban.
Khosa, G. (2010). The South African national education evaluation system: What will make it work? Occasional paper 1.

Koornhof, H. (2012). From conception to consumption: an examination of the intellectual process of producing textbooks for Foundation Phase in South Africa. (M Ed dissertation, University of the Witwatersrand).

Kotze, J., Fleisch, B. & Taylor S. (2018). Alternative forms of early grade instructional coaching: Emerging evidence from field experiments in South Africa. Paper presented at the REAL conference, Cambridge University.

Kraft, M. A., Blazar, D. & Hogan, D. (2016). The effect of teacher coaching on instruction and achievement: A meta-analysis of the causal evidence. (Brown University Working Paper.*)*

Leithwood, K., Jantzi, D., Earl, L., Watson, N., Levin, B. & Fullan, M. (2004). Strategic leadership for large-scale reform: the case of England's national literacy and numeracy strategy. *School Leadership & Management*, *24*(1): 57–79.

Levin, B. & Fullan, M. (2008). Learning about system renewal. *Educational Management Administration & Leadership*, *36*(2): 289–303.

Lim, L. & Apple, M. W. (eds). (2016). *The strong state and curriculum reform: Assessing the politics and possibilities of educational change in Asia*. London: Routledge.

Macdonald, C. A. (2002). Are the children still swimming up the waterfall? A look at literacy development in the new curriculum. *Language Matters: Studies in the Languages of Southern Africa*, *33*(1): 111–141.

Mackie, J. M. (2007). Beyond Learning to Read: An Evaluation of a Short Reading Intervention in the Ilembe District of Kwazulu-Natal. (M Ed Research Report, University of Kwazulu-Natal).

Malda, M., Nel, C. & Van de Vijver, F. J. (2014). The road to reading for South African learners: The role of orthographic depth. *Learning and Individual Differences*, *30*: 34–45.

Maphumulo, T. (2010). An Exploration of how Grade 1 IsiZulu Teacher's Teach Reading. (M Ed, University of Kwazulu-Natal).

Masino, S., & Niño-Zarazúa, M. (2016). What works to improve the quality of student learning in developing countries?. *International Journal of Educational Development*, *48*: 53–65.

Masterson, L. (2013). A case study of Foundation Phase teachers' experiences of literacy coaching in the GPLMS programme. (M Ed Research Report, University of the Witwatersrand).

Maswanganye, B. (2010). The Teaching of First Additional Language Reading in the Grade 4 in Selected Schools in the Moretele Area Project Office. (M Ed, UNISA).

Mbatha, T. (2012). IsiZulu teachers' perceptions of the foundations for learning literacy programme in four township primary schools. *Per Linguam: a Journal of Language Learning Per Linguam: Tydskrif vir Taalaanleer*, *28*(1): 59–72.

McEwan, P. J. (2015). Improving learning in primary schools of developing countries: A meta-analysis of randomized experiments. *Review of Educational Research*, *85*(3): 353–394.

Mensah, F. J. (2017).The experiences of primary school educators with the national reading strategy in Mbabane Circuit, iNgwavuma. (Doctoral dissertation, University of Zululand).

Mourshed, M., Chijioke, C. & Barber, M. (2010). *How the world's most improved school systems keep getting better*. McKinsey & Company.

Murahwa, S. (2009). Combining 'Phonics' and 'Whole Language' Approaches in Teaching Reading: A Case of Phenduka Literacy Project in an Alexandra School. (MA Research Report, University of the Witwatersrand).

Muralidharan, K. & Sundararaman, V. (2013). *Contract teachers: Experimental evidence from India* (No. w19440). National Bureau of Economic Research.

Nadel, S., & Pritchett, L. (2016). Searching for the Devil in the Details: Learning about Development Program Design. Working Paper for Centre for Global Development.

NEEDU. (2012). *The state of literacy teaching and learning in the foundation phase*. Pretoria: Ministry of Education.

NEEDU. (2013). *Teaching and learning in rural primary schools*. Pretoria: Ministry of Education.

O'Day, J. A. & Smith, M. S. (1993). Systemic reform and educational opportunity. In S. Fuhrman (ed) *Designing coherent education policy: Improving the system*, San Francisco: Jossey-Bass.

Oketch, M., Ngware, M., Mutisya, M., Kassahun, A., Abuya, B. & Musyoka, P. (2014). When to randomize: Lessons from independent impact evaluation of reading to learn (RtL) programme to improve literacy and numeracy in Kenya and Uganda. *Peabody Journal of Education*, *89*(1): 17–42.

Perry, K. (2009). Languages, literacies and libraries: a view from Africa. In J. A. Kleifgen & G. Bond (eds) *The Languages of Africa and the Diaspora.* Bristol, UK; Buffalo, NY: Multilingual Matters.

Piper, B. (forthcoming) Effectiveness of Teachers' Guides in Developing Countries: Scripting, learning outcomes, and classroom utilization.

Piper, B. (forthcoming) Scaling up successfully? Lessons from Kenya's Tusome national literacy program for us.

Piper, B. (2009). Integrated Education Program: Impact study of SMRS using Early Grade Reading Assessment in three provinces in South Africa. International Research Triangle Park, North Carolina: USAID/Southern Africa.

Piper, B., Schroeder, L. & Trudell, B. (2015). Oral reading fluency and comprehension in Kenya: Reading acquisition in a multilingual environment. *Journal of Research in Reading, 39*(2): 133–152.

Piper, B., Zuilkowski, S. S. & Mugenda, A. (2014). Improving reading outcomes in Kenya: First-year effects of the PRIMR Initiative. *International Journal of Educational Development*, *37*: 11–21.

Piper, B. & Zuilkowski, S. S. (2015a). Teacher coaching in Kenya: Examining instructional support in public and nonformal schools. *Teaching and Teacher Education*, *47*: 173–183.

Piper, B. & Zuilkowski, S. S. (2015b). Assessing reading fluency in Kenya: Oral or silent assessment? *International Review of Education*, *61*(2): 153–171.

Piper, B. & Zuilkowski, S. S. (2016). The role of timing in assessing oral reading fluency and comprehension in Kenya. *Language Testing*, *33*(1): 75–98.

Piper, B., Zuilkowski, S. S., Kwayumba, D. & Strigel, C. (2016). Does technology improve reading outcomes? Comparing the effectiveness and cost-effectiveness of ICT interventions for early grade reading in Kenya. *International Journal of Educational Development*, *49*: 204–214.

Popova, A., Evans, D. & Arancibia, V. (2016). Training teachers on the job: what works and how to measure it (English). Policy Research working paper; no. WPS 7834; WDR 2018 background paper. Washington, D.C.: World Bank Group. Retrieved from http://documents.worldbank.org/curated/en/122031474915232609/Training-teachers-on-the-job-what-works-and-how-to-measure-it. (Accessed 26 June 2018).

Pretorius, E. (2015). Failure to launch: Matching language policy with literacy accomplishment in South African schools. *International Journal of the Sociology of Language*, *2015*(234): 47–76.

Pretorius, E. J. & Currin, S. (2010). Do the rich get richer and the poor poorer? The effects of an intervention programme on reading in the home and school language in a high poverty multilingual context. *International Journal of Educational Development*, *30*(1): 67–76.

Pretorius, E. J. & Machet, M. P. (2004). The socio-educational context of literacy accomplishment in disadvantaged schools: Lessons for reading in the early primary school years. *Journal of Language Teaching*, *33*: 45–60.

Pretorius, E. J. & Mokhwesana, M. M. (2009). Putting reading in Northern Sotho on track in the early years: Changing resources, expectations and practices in a high poverty school. *South African Journal of African Languages*, *29*(1): 54–73.

Pretorius, E. J. & Spaull, N. (2016). Exploring relationships between oral reading fluency and reading comprehension amongst English second language readers in South Africa. *Reading and Writing*, *29*(7): 1449–1471.

Pritchett, L. & Beatty, A. (2015). Slow down, you're going too fast: Matching curricula to student skill levels. *International Journal of Educational Development*, *40*: 276–288.

Ralaingita, W. & Wetterberg, A. (2011). Gauging Program Effectiveness with EGRA: Impact

Evaluations in South Africa and Mali. *The Early Grade Reading Assessment: Applications and interventions to improve basic literacy.* A. W. Gove, A. Research Triangle Park, NC, RTI Press.
Raudenbush, S. W. (2005). Learning from attempts to improve schooling: The contribution of methodological diversity. *Educational Researcher*, *34*(5): 25–31.
Raudenbush, S. W. (2008). Advancing educational policy by advancing research on instruction. *American Educational Research Journal*, *45*(1): 206–230.
Reeves, C., Heugh, K., Prinsloo, C H., Macdonald, C., Netshitangani, T., Alidou, H., Dierdericks, G. & Herbst, D. (2008). Evaluation of Literacy Teaching in Primary Schools in Limpopo. Polokwane: Limpopo Department of Education.
Reeves, J. (2011). Teacher learning by script. *Language Teaching Research* (2010), *14*(3): 241–258.
Rincón-Gallardo, S. (2016). Large scale pedagogical transformation as widespread cultural change in Mexican public schools. *Journal of Educational Change*, *17*(4): 411–436.
Rodseth, V. (2002). Developing main language instruction: development of The Molteno Project—literacy, language and educator development. *Perspectives in Education*, *20*(1): 97–109.
Sahlberg, P. (2014). *Finnish lessons 2.0: What can the world learn from educational change in Finland?* New York, NY: Teachers College Press, Columbia University.
Sahlberg, P. (2016). The global educational reform movement and its impact on schooling. *The Handbook of Global Education Policy*, 128–144. Hoboken, New Jersey: Wiley-Blackwell.
Sailor, M., Hoffman, J. & Matthee, B. (2007). South African schools that promote literacy learning with students from low-income communities. *Reading Research Quarterly*, *42*(3): 354–387.
Sailor, M., Hoffman, J., Pearson, P., Beretvas, S. & Matthee, B. (2010). The effects of first- and second-language instruction in rural South African schools. *Bilingual Research Journal*, *33*: 21–41.
Sayed, Y., Alhawsawi, S., Mwale, S. & Van Niekerk, R. (2013). The state and education policy change in South Africa: Walking the tightrope between choice and equity. In D. Turner & H. Yolcu (eds) *Neo-liberal educational reforms: A critical analysis*. London: Routledge.
Schollar, E. (2001). A review of two evaluations of the application of the READ primary schools program in the Eastern Cape Province of South Africa. *International Journal of Educational Research, 35*(2): 205–216.
Schollar, E. (2005). The Evaluation of the Learning for Living Project: Final Report, Business Trust and the READ Organisation.
Shalem, Y. (2017). Scripted lesson plans. In B. Barrett, U. Hoadley, and J. Morgan (eds) *Knowledge, Curriculum and Equity: Social Realist Perspectives*. London: Routledge.
Shalem, Y. & Slonimsky, L. (2010). Seeing epistemic order: Construction and transmission of evaluative criteria. *British Journal of Sociology of Education*, *31*(6): 755–778.
Shalem, Y., Steinberg, C., Koornhof, H. & De Clercq, F. (2016). The what and how in scripted lesson plans: the case of the Gauteng Primary Language and Mathematics Strategy. *Journal of Education*, (66): 13–36.
Snilstveit, B., Stevenson, J., Menon, R., Phillips, D., Gallagher, E., Geleen, M., Jobse, H., Schmidt, T. & Jimenez, E. (2016). The impact of education programmes on learning and school participation in low-and middle-income countries. International Initiative for Impact Evaluation (3ie), a systematic review.
Snow, C. E. (2017). Early literacy development and instruction: An overview. In N. Kucirkova, C. E. Snow, V. Grøver & C. McBride-Chang (eds) *The Routledge International Handbook of Early Literacy Education: A Contemporary Guide to Literacy Teaching and Interventions in a Global Context*, 5–13. New York: Routledge.
Spaull, N. (2013). Poverty & privilege: Primary school inequality in South Africa. *International Journal of Educational Development*, *33*(5): 436–447.
Spaull, N. (2015). Examining oral reading fluency among rural Grade 5 English Second Language (ESL) learners in South Africa: An analysis of NEEDU 2013. *South African Journal of Childhood Education*, *5*(2) 44–77.

Spaull, N. (2016). *Excessive class sizes in the Foundation Phase RESEP working paper*. http://resep.sun.ac.za/wp-content/uploads/2016/05/RESEP-Policy-Briefs_ClassSize-EMAIL.pdf. (Accessed 12 September 2017).

Spaull, N., Pretorius, L., & Mohohlwane, N. (2017). Investigating the Comprehension Iceberg: Developing empirical benchmarks for early grade reading in agglutinating African languages. (Working paper).

Spaull, N., Van der Berg, S., Wills, G., Gustafsson, M. & Kotzé, J. (2016). Laying Firm foundations: Getting reading right. RESEP, Stellenbosch University.

Stein, P. & Slonimsky, L. (2006). An eye on the text and an eye on the future: Multimodal literacy in three Johannesburg families. In K. Pahl & J. Rowsell (eds) *Travel notes from the new literacy studies: Instances of practice* (Vol. 4). Clevedon: Multilingual Matters.

Stone-Johnson, C. (2016). *Generational identity, educational change, and school leadership*. New York: Routledge.

Taylor, N. (2011). The National School Effectiveness Study (NSES): Summary for the synthesis report. Johannesburg: JET education services.

Taylor, N. & Moyane, J. (2005). Khanyisa Education Support Programme: Baseline study part 1: communities, schools and classrooms. Johannesburg: JET Education Services.

Taylor, N., Van der Berg, S. & Mabogoane, T. (2013). *Creating effective schools*. Cape Town: Pearson Education South Africa.

Taylor, S. (2017). How context mediates the effectiveness of teacher support interventions: Evidence from a multi-arm randomised experiment in South Africa. Paper presented in ESSA.

Thaler, R. H. & Sunstein, C. R. (1999). *Nudge: Improving decisions about health, wealth, and happiness*. New Haven: CT Yale University Press.

The Economist. (2017). [online] Available at https://www.economist.com. (Accessed 22 June 2018).

Turnbull, M. (2010). Self-evaluation a means of raising achievement/impossible to implement? Paper presented to Commonwealth Council for Educational Administration and Management (CCEAM) in South Africa.

Van der Berg, S. (2009). *The persistence of inequalities in education*. Oxford: Oxford University Press.

Van der Berg, S. (2015). What the Annual National Assessments can tell us about learning deficits over the education system and the school career year. RESEP, Stellenbosch University.

Van der Berg, S., Spaull, N., Wills, G., Gustafsson, M. & Kotzé, J. (2016). Identifying Binding Constraints in Education. RESEP Stellenbosch University.

Van der Berg, S., Taylor, S., Gustafsson, M., Spaull, N. & Armstrong, P. (2011). Improving education quality in South Africa. Pretoria: Report for the National Planning Commission.

Vella, S., Schwartländer, B., Sow, S. P., Eholie, S. P. & Murphy, R. L. (2012). The history of antiretroviral therapy and of its implementation in resource-limited areas of the world. *Aids*, *26*(10): 1231–1241.

Venkat, H. (2013). Using temporal range to theorize early number teaching in South Africa. *For the Learning of Mathematics*, *33*(2): 31–37.

Veriava, F. (2017). Basic Education Rights Handbook. Funding Basic Education. http://section27.org.za/wp-content/uploads/2017/02/Chapter-2.pdf. (Accessed 7 August 2017).

Von Holdt, K. (2010). Nationalism, bureaucracy and the developmental state: the South African case. *South African Review of Sociology*, *41*(1): 4–27.

Wagner, D. A. (2003). Smaller, quicker, cheaper: Alternative strategies for literacy assessment in the UN Literacy Decade. *International Journal of Educational Research*, *39*(3): 293–309.

Western Cape Education Department (2006) LITERACY AND NUMERACY STRATEGY 2006 – 2016 https://wcedonline.westerncape.gov.za/documents/literacy_numeracy_strategy/e-LitNumStrat.pdf. (Accessed 26 June 2018).

World Bank. (2018). World Bank Development Report 2018. *Learning to Realize Education's Promise*. Washington DC.

Index

Page numbers in *italics* indicate figures and tables. The letter n after a page number indicates a footnote.