EARLY GRADE READING AND MATHEMATICS INTERVENTIONS IN SOUTH AFRICA

INTERVENTIONS

EDITED BY

Nic Spaull Stephen Taylor

OXFORD

EARLY GRADE READING AND MATHEMATICS
INTERVENTIONS IN SOUTH AFRICA

INTERVENTIONS

EDITED BY
Nic Spaull Stephen Taylor

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide. Oxford is a registered trade mark of
Oxford University Press in the UK and in certain other countries.

Published in South Africa by
Oxford University Press Southern Africa (Pty) Limited

Vasco Boulevard, Goodwood, N1 City, Cape Town, South Africa, 7460
P O Box 12119, N1 City, Cape Town, South Africa, 7463

First published 2022

Early Grade Reading and Mathematics Interventions in South Africa

Print ISBN 978 0 19 074527 1
ePDF ISBN 978 0 19 074119 8

First impression 2022

Typeset in Tisa Pro Light 10pt on 12pt

Acknowledgements
Publisher: Megan Hall
Managing Editor: Ashly Fraser
Editor: Lorna Morris
Proofreader: Lindsey Morton
Indexer: Tanya Paulse
Designer: Judith Cross-McCulloch
Typesetter: JSquare Studio

table of contents

table of contents

continued

editors

Nic Spaull is an associate professor in the Research on Socioeconomic Policy (ReSEP) group within the Department of Economics at Stellenbosch University. He has published extensively on the inequality of educational opportunity in South Africa, focusing on assessment, education policy and early grade reading and mathematics. He is a P-rated NRF researcher and has held visiting fellowships at Stanford University in the United States and at the Organisation for Economic Cooperation and Development in Paris. Prof. Spaull has been involved in a number of research projects for local and international organisations, including the South African Presidency, the Department of Basic Education, UNICEF, the European Union, UNESCO, ESRC, and SACMEQ. He is Special Adviser to the Allan & Gill Gray Philanthropies and also advises numerous NGOs, policy-makers, and grant-making bodies.

Stephen Taylor is Director of Research, Monitoring and Evaluation in the Department of Basic Education, South Africa. Since 2007, he has been doing research on education in South Africa and the surrounding region, with a focus on the links between education and economic development. He has been a principal investigator on several large scale evaluations, including the Early Grade Reading Study. His academic work focuses on impact evaluation of education interventions, measuring educational performance, and equity in educational outcomes. He holds a PhD in economics from the University of Stellenbosch.

contributors

Kaitlin Alsofrom is a PhD candidate at the University of Witwatersrand and the lead at Class Act Educational Services.

Cally Ardington is a professor in the Southern Africa Labour and Development Research Unit in the School of Economics at the University of Cape Town.

Jayesh Bhana is an independent education consultant working in the area of human capacity development as a facilitator of learning.

Megan Borole is the director of Firdale Consulting.

Aarnout Brombacher is Director of the NumberSense Mathematics Programme (South Africa) and the team leader on the National Numeracy Programme (Malawi).

Dhianaraj Chetty is Manager: Monitoring and Evaluation at the National Education Collaboration Trust (NECT).

Brahm Fleisch is a professor in the division of Educational Leadership and Policy Studies at the School of Education at the University of the Witwatersrand.

Vanessa Francis is the country director at Room to Read South Africa.

Mellony Graven is a full professor at Rhodes University and holds the South African Numeracy Chair.

Penelope Groome is an executive director at Class Act Educational Services and is the materials production manager at the National Education Collaboration Trust.

Nicola Harris is the founder and chief executive officer of Click Learning.

Eleanor Hazell is a research associate in the Department of Childhood Education at the University of Johannesburg and Executive Manager: Monitoring and Evaluation at JET Education Services.

Permie Isaac is the head of content at Funda Wande.

Anil Kanjee is a research professor and head of the Assessment for Learning Research Niche Area at the Tshwane University of Technology.

Jenny Katz is the research and development manager at the Molteno Institute for Language and Literacies NPC.

Cally Kuhne is an independent early learning consultant who until recently was a senior education specialist at the University of Cape Town's Schools Development Unit and a lead writer/project manager of the R-Maths Programme.

Vuyokazi Mafilika is a maths specialist at Funda Wande in the Eastern Cape (Buffalo City Metro and Sarah Baartman Districts).

Nwabisa Makaluza is the head of the Limpopo Intervention at Funda Wande.

Bruce McDougall is an analyst at Firdale Consulting.

Veronica McKay is a professor extraordinarius in the College of Education at the University of South Africa.

Qetelo Moloi is an associate researcher in the Department of Childhood Education at the University of Johannesburg.

Bokang Mpeta was formerly the project manager of the Teaching Assistant Arm of the Limpopo Intervention at Funda Wande.

Sihlobosenkosi Mpofu is a senior mathematics specialist at Bala Wande.

Thobeka Ndamase is a teacher by profession and a maths specialist at Funda Wande in the Eastern Cape (Nelson Mandela Metro District).

Catherine Ngwane is the programme operations director at Room to Read South Africa.

Joy Olivier is an independent research consultant.

Siân Rees is the senior research and development officer at the Molteno Institute for Language and Literacies NPC.

Nicky Roberts is a professor in mathematics education at the University of Fort Hare.

Lorna Sako is a maths specialist based in Limpopo Province, supporting the implematation of the Bala Wande strategy in participating Sepedi-LOLT schools.

Ingrid Sapire is based at the University of Witwatersrand and is the head of mathematics content for Bala Wande.

Mmapula Seolana is a literacy facilitator for Maths and Sepedi at Funda Wande in Limpopo.

Garth Spencer-Smith is a research associate in the Department of Childhood Education at the University of Johannesburg.

Jabu Thomo is a Jumpstart Programmes director.

Kerri Thomsen is the associate director of literacy at Room to Read.

John Thurlow is the Reading Programmes manager at the National Education Collaboration Trust.

Hamsa Venkat holds the Naughton Chair in Early Years/Primary STEM Education at Dublin City University and a visiting professorship at the University of the Witwatersrand.

foreword

How can we improve literacy and numeracy outcomes for children? Study after study—in South Africa and in many other countries—demonstrates two key facts about literacy and numeracy skills. First, children who master literacy and numeracy skills in the early years continue to succeed later in their school careers. Adults who have mastered these skills while in school have better work and life outcomes. In case there were any doubt, these skills matter enormously. Second, far too few students master these skills in the early grades. Children can attend years of school and still not be able to read and write or do mathematics at levels that help them to live their richest, fullest possible lives. But identifying that problem—while crucial—is much easier than proposing a solution. Implementing, testing, and scaling a solution that has proven effective are harder yet.

In South Africa, the last decade has seen improvements in measures of student learning, as documented by Spaull and Taylor in the first chapter of this volume. In the same decade, a plethora of programmes have sought to improve these outcomes, including workbooks, teacher coaching, teaching assistants, student readers in multiple languages, and many more. Several of these programmes have been evaluated carefully to identify a causal impact. This volume documents a wide range of these experiences, from the nuts and bolts of implementation to what has worked and what has not to boost student learning.

From my perspective, South Africa has one of the richest collections of educational innovation, experimentation, and complementary research to evaluate impact across the African continent and even across all low- and middle-income countries. Furthermore, as you'll see in this volume, there is a rich interplay between government implementation, non-government implementation, and government and academic researchers that forms a model for other countries. I've already seen South Africa's innovation and research on coaching influencing global conversations on how to improve education. This volume reveals how much more the world has to learn from South Africa's ongoing work.

Readers can expect three major contributions from this volume. First, for readers who want to know 'what works', several chapters document the impact of effective (and ineffective) interventions to boost learning. This comes in two forms. Some chapters provide novel estimates of programme impacts; others yield new perspectives by synthesising and contextualising existing impact estimates. Second, readers will gain insight into why we observe the impacts that we do. In-depth qualitative studies and institutional histories complement quantitative studies to yield new perspectives. Third, readers will find rich implementation data. An enduring challenge of impact evaluation studies is that they often tell you that a programme worked without telling you very much about the programme itself. In this volume, contributors provide rich implementation detail that codifies institutional knowledge both for policymakers in South Africa and for those in other countries who may want to replicate these experiences.

South Africa—just like most other countries in the world—still has a long way to go in improving the quality of its education system. But improvements in the last decade are encouraging, as are positive impacts from several of the programmes you'll read about in this volume. As in other contexts at all income levels, scaling programmes while retaining an impact and managing costs is a challenge. But readers in South Africa and in other countries will see both promising directions—especially teacher coaching and providing teachers' assistants—and an example of a dynamic process of experimentation and rigorous evaluation that will be crucial to any education system seeking to provide quality education to its next generation.

David K. Evans
Center for Global Development

preface

This edited volume is one of three books in a series focusing on developments in early-grade reading and mathematics in South Africa between 2010 and 2022. The first volume is *Early Grade Reading in South Africa*, edited by Nic Spaull and Elizabeth Pretorius, the second volume is *Early Grade Mathematics in South Africa*, edited by Hamsa Venkat and Nicky Roberts, and the third is *Early Grade Reading and Mathematics Interventions in South Africa*, edited by Nic Spaull and Stephen Taylor. Collectively the three books bring together 77 authors from disciplines including economics, linguistics, literacy studies, mathematics education, teacher education, and policy studies. Although their domains and methods of analysis may differ, all authors grappled with the same underlying question: why is it that so few young children in South Africa acquire the building blocks of reading and mathematics in the first years of school? While international large-scale assessments have drawn increasing attention to learning outcomes at the primary school level, there is now a broad-based consensus that the roots of the problem lie even earlier than upper primary school. International assessments like PIRLS and TIMSS show that 60–80% of Grade 4 and 5 learners cannot read for meaning or calculate using the four operations, but emerging research documented in these volumes highlights that more than 50% of learners at the end of Grade 1 do not know all the letters of the alphabet, and cannot add and subtract single-digit numbers.

It is this challenge that animates the research across these three volumes, with an analytic focus on lessons learned in the last decade (2010–2022). While learning outcomes in South Africa before the Covid-19 pandemic were improving quickly by international standards, the chapters included here present evidence for both optimism and alarm. Optimism because system-wide improvements do not happen accidentally or in a vacuum. Alarm because in 2022 it is still the case that the dignity and life-chances of millions of children in South Africa are foreclosed because they do not learn to read for meaning, or do mathematics with understanding in the first three years of school.

As a group of scholars committed to understanding and documenting the roots of both blockages and breakthroughs in reading and mathematics, it is our hope that you, the reader, find this new research interesting, helpful, generative and challenging.

Acknowledgements

The editors thank the Allan & Gill Gray Philanthropies for their financial support of two in-person authors’ workshops, and for funding the excellent editorial assistant, Jess Qvist. We also thank Oxford University Press SA, who kindly agreed to bear the editorial costs of publishing these three books in an Open Access format, making them available to a wider audience than would otherwise be the case.

Statement of peer review

All chapters in this book have been peer-reviewed, with the overall process managed by Oxford University Press SA. In a double-blind process, the reviews were anonymous and neither authors nor reviewers knew the identity of the reviewers or authors. The entire book’s manuscript was also peer-reviewed by an international expert in the field.

abbreviations and key terms

ANAs	Annual National Assessments
CAPS	Curriculum and Assessment Policy Statement
CASME	Centre for the Advancement of Science and Mathematics
DBE	Department of Basic Education (after 2009)
DHET	Department of Higher Education and Training (after 2009)
DOE	Department of Education (before 2009)
ECDOE	Eastern Cape Department of Education
EFAL	English as a First Additional Language, according to the South African curriculum
EGM	Early Grades Mathematics
EGRA	Early Grade Reading Assessment
EGRS	Early Grade Reading Study
ePIRLS	electronic-Progress in International Reading Literacy Study
FAL	First Additional Language
FP	Foundation Phase (Grades 1 to 3)
GDE	Gauteng Department of Education
GPLMS	Gauteng Primary Language and Mathematics Strategy
HL	Home Language
HOD	Head of Department
ICT	Information and Communications Technology
IIAL	Incremental Introduction to African Languages
IP	Intermediate Phase (Grades 4 to 6)
learner	a child enrolled at school, or someone learning a language (whether a child or an adult)
LAB	learner activity book
LOLT	Language of Learning and Teaching
LPDOE	Limpopo Department of Education
LTSM	Learning and Teaching Support Materials
MSAP	Mental Starters Assessment Project
NCS	National Curriculum Statement
NECT	National Education Collaboration Trust
NEIMS	National Education Infrastructure Management System
NGO	non-governmental organisation
PANSALB	Pan South African Language Board
PIRLS	Progress in International Reading Literacy Study
PSRIP	Primary School Reading Improvement Programme
RCTs	randomised control trials
R-Maths	Grade R Mathematics Project
SA	Subject Advisor
SACMEQ	Southern Africa Consortium for Monitoring Educational Quality, *or* the Southern and Eastern Africa Consortium for Monitoring Educational Quality
SMRS	Systematic Method for Reading Success
SMS	School Monitoring Survey
SMT	school management teams
SONA	State of the Nation Address
TA	Teaching Assistant
TaRL	Teaching at the right level
TG	Teacher guide
TIMSS	Trends in International Mathematics and Science Study
UCT	University of Cape Town
WCED	Western Cape Education Department

01

Impact or scale? The trade-offs of early grade reading and mathematics interventions in South Africa

NIC SPAULL & STEPHEN TAYLOR

Abstract

In this chapter we document the lineage of early grade reading and mathematics interventions in South Africa from 2010 to 2022. We develop an intervention typology using four dimensions: (1) programme dosage, (2) implementation scale, (3) government buy-in, and (4) evidence of impact and evaluation method. Using this rubric we review 14 different interventions and find three distinct families or intervention approaches. The first is that of light-touch large-scale interventions that were implemented across multiple provinces and at scale (4,000+ schools) but lack causal evaluations of impact. The second category is province-level interventions which have quasi-experimental evidence of some impact. The final group are small scale NGO-led interventions, some of which show causal evidence of impact on reading and mathematics, but all of which operate in 50 schools or fewer. We find a clear trade-off between dosage and scale across the programmes and contrast the two approaches (1) large-scale, trickle-down, train-the-trainer approaches, with (2) small-scale, bottom-up, in-classroom support models. The latter typically have a coach-to-teacher ratio of 1:30. In a new analysis we show that the ratio of subject advisors to teachers in the 'train-the-trainer' model is exceedingly large (1:500) in five of the nine provinces, making meaningful support impossible. However, even with intensive in-

KEYWORDS
early grade reading and mathematics, evaluation, interventions, subject advisors

classroom support, gains in reading and mathematics outcomes are often lower than one would hope. We hypothesise that as a country we may be reaching a 'low ceiling' beyond which it is not possible to extend unless the superstructures within which these interventions are implemented are reformed. These superstructures include teacher knowledge, accountability relations, class sizes, and learner poverty.

1 Introduction: The 'long' decade from 2009-2022

The 'long' decade from 2009 to 2022 in South Africa has been one of both promise and peril. The first half of the decade saw the introduction of a new curriculum, a new assessment system, and a new universal set of workbooks. It was also the first time that international assessments began to show improvements in learner outcomes. The first of these was the 2011 Trends in International Mathematics and Science Study (TIMSS) which showed national improvement since 2003. However, one swallow does not make a summer and these results were seen as tentative, at least until subsequent assessments (SACMEQ[1] 2013, TIMSS 2015 and PIRLS 2016) all showed similar gains compared to their previous rounds of assessment. Unfortunately, the latter part of this long decade (2020-2022) has been dominated by the largest single shock in a generation: the Covid-19 pandemic and the social, economic, and educational devastation left in its wake. Assessments across multiple grades and provinces have revealed learning losses of approximately one full year of school with the average South African ten-year-old in 2022 knowing less than the average nine-year-old in 2019 (Ardington et al. 2021; Shepherd & Mohohlwane 2021; Van der Berg et al. 2022a).

1.1 A long decade of progress (2009-2022)

Yet notwithstanding the above, most of the decade can be seen as one of progress and maturation, at least on the education front. The decade began with the splitting of the single department of education in 2010 into one governing schools (the Department of Basic Education, DBE) and one governing universities and colleges (the Department of Higher Education and Training, DHET). At the same time the new and now longest-serving Minister of Basic Education, Ms Angie Motshekga, was appointed with an explicit mandate to introduce a new curriculum (Curriculum Assessment Policy Statement, CAPS), a new universal learner-level workbook programme (the DBE Workbooks) and a new universal primary school assessment system (the Annual National Assessments, ANAs). These three reforms are discussed in greater detail in the current volume (McKay & Spaull 2022), as well as in the companion volume on reading (Pretorius et al. 2022), and mathematics (Nuga Deliwe & Van der Berg 2022).

In addition to these far-reaching reforms of curriculum, materials, and assessment, in 2009 the DBE also universalised an existing school feeding scheme, the National School Nutrition Programme (NSNP) by extending it to no-fee high schools (JET 2016: 2). This programme now reaches more than 9-million children every day accounting for

1 SACMEQ is the Southern and Eastern African Consortium for Monitoring Educational Quality (SACMEQ) and PIRLS is the Progress in International Reading Literacy Study.

all children in no-fee schools and approximately 75% of all children in the country. With respect to early grade reading and mathematics there has also been considerable policy innovation with the recent publication of the Framework for Teaching Mathematics for Understanding (DBE, 2018) and the National Framework for the Teaching of Reading in African Languages in the Foundation Phase (DBE 2020a).

Perhaps most importantly, the period 2011 to 2019 saw an improvement in educational outcomes as reported in a number of independently administered and nationally-representative surveys of educational achievement. This includes statistically significant improvements in both reading (PIRLS) and mathematics (TIMSS). The PIRLS studies show that the percentage[2] of Grade-4 learners who can read for meaning rose from 13% (2006) to 18% (2011) to 22% (2016) (Van Staden & Gustafsson, 2022). For mathematics the TIMSS studies show that at the Grade 9 level the percentage of South African learners who could reach the low international benchmark rose from 10% (2003) to 24% (2011) to 34% (2015) and finally to 41% (2019)[3], although it should be noted that the TIMSS-Numeracy assessment at the Grade 5 level shows no improvement between 2015 and 2019 (Spaull et al. 2022a).

For the purposes of the current volume, and this chapter, it is of special relevance to note that the long decade 2009-2022 has also seen a rapid expansion in the number of interventions targeting early grade reading and mathematics, a refinement in the design and quality of those interventions and evaluations, and greater public awareness of the importance of foundational skills. This is the focus of our current volume and the topic to which we turn after first discussing some of the persistent challenges that have plagued not only South African education but also educational interventions and attempts at improving outcomes.

1.2 A long decade of unaddressed challenges (2009-2022)

In addition to the hopeful signs of progress listed above, there have also been a number of interventions aimed at improving reading and mathematics outcomes in the Foundation Phase, some of which have shown measurable gains, albeit somewhat more modest than one might hope. One possible explanation behind these modest gains is our ‘low ceiling’ hypothesis where interventions, even very successful interventions, reach a low ceiling beyond which they are unable to extend. This may be because the superstructures within which these interventions are implemented have remained largely unchanged. These superstructures include teacher knowledge, accountability relations, class sizes, and learner poverty.

The low level of teacher content knowledge is perhaps the most obvious binding constraint across interventions and one that has not changed much in the last decade. The only two studies to test a nationally representative sample of primary school teachers on a content knowledge test in South Africa were the SACMEQ 2007 and SACMEQ 2013 studies. The former showed that 79% of Grade 6 mathematics teachers could not score 60% on a primary school mathematics test (Venkat & Spaull 2015). The more recent

2 For further references for these figures see Howie et al. (2008, 26) for the 13% in 2006, and DBE (2020b, 69) for the 18% in 2011 and the 22% in 2016 (DBE, 2020b, 69).

3 See Reddy et al. 2003, 25; Mullis et al. 2012, 115; Zuze et al. 2018, 23; & Reddy et al. 2020, 5 respectively.

SACMEQ 2013 study revealed that only 41% of South African Grade 6 mathematics teachers were rated as having 'good proficiency in mathematics' compared to 95% of teachers in Kenya and 87% of teachers in Zimbabwe (Awich 2021, 62).

While it is true that newer teachers entering the profession have higher levels of content knowledge than their more senior colleagues (Armstrong 2015, 136), even these younger and better-performing teachers still have exceedingly low levels of content knowledge. This is most likely because teaching still has the lowest entry requirements of all university degrees. New research shows that across all incoming Bachelor of Education students across the country, only one in five scored 50% or higher for mathematics in matric, compared to half (54%) of incoming students for other degrees (Van der Berg et al. 2022, 102). At the largest single university (UNISA) only one in ten incoming Bachelor of Education students scored more than 50% for mathematics in matric (ibid). In 2018, Bowie et al. (2019) tested a sample of 488 first-year Bachelor of Education (BEd) students and 282 final year BEd students from three typical South African universities. This was on a test that included only primary-school level items (43 questions). Despite this, the average score for first-year BEds was 52% and only 54% for final-year BEd students on the same test, despite four additional years of full-time study. Taylor & Mawoyo (2022) & Taylor (2021) discuss these initial teacher-education challenges in more detail, yet it is sufficient to say that at least half of South African primary school teachers lack the primary-school-level content knowledge needed to teach their subjects.

In the current volume Chetty et al. report on an intervention targeting Foundation Phase teachers as well as their subject advisors. After testing more than 1 000 Foundation Phase teachers across multiple provinces, and finding some improvement in test scores following training, the average post-improvement score was still only 49%. Subject Advisors scored slightly higher on the test, but only marginally so in most instances. This was notwithstanding the fact that the test assessed only basic knowledge about the policy requirements of the curriculum and teaching reading.

Other superstructures that remain unchanged include accountability relationships between schools and districts as well as between teachers and principals. Following union objections to the use of ANAs as an accountability tool, the only national exam at primary school level was discontinued in 2014 after being introduced in 2011 (Nuga Deliwe & Van der Berg 2022). Relatedly, in 2016 the Volmink Commission of Inquiry into 'Jobs for Cash' (DBE, 2016) revealed that "corruption is endemic to greater and lesser degrees in the entire education system" pointing to undue union influence. And further that "of the 81 cases that were investigated, 38 cases provided grounds for either reasonable suspicion requiring further investigation or point directly to wrongdoing amounting to criminal conduct" (DBE 2016, 18). As of July 2022, no government official implicated in the scandal has been prosecuted or suspended.

Finally, one of the perennial concerns in South Africa is school-level resources and funding. Previous research has shown an overall decline in real per-learner funding of −2.3% between 2009 and 2018 (Spaull et al. 2020), primarily due to above-inflation wage increases for teachers accompanied by a rise in births in 2005. That study argues that "provinces seem to be 'coping' by implementing cost-saving measures such as hiring freezes and leaving vacant posts unfilled" (Spaull et al. 2020, 3). This seems to be borne out in the current set of volumes where Spaull et al. (2022a) using TIMSS-Grade 5 data show large increases in average class sizes between 2015 and 2019. This study identified significant increases in the percentage of Grade 5 children in very large class sizes (>50), doubling from 16% (2015) to 34% (2019). In

addition to the negative impacts of larger classes, these class-size increases should also be seen as a proxy for a squeeze on school-level resources more generally.
Over and above the systemic issues listed above, the single biggest factor influencing education in South Africa now is the long tail of the Covid-19 pandemic and consequences of school closures and learning losses. In the middle of the pandemic Ardington et al. (2021) estimated learning losses of approximately 70% of a year of learning. More recently, and using data from more than 80,000 learners in the Western Cape, Van der Berg et al. (2022b) estimate losses at the primary level to be approximately one full year of learning, with larger losses in mathematics and at earlier grades.

1.3 The long decade of early grade reading and mathematics interventions

The purpose of the present chapter, and of this volume more generally, is to document and synthesise the findings from early grade reading and mathematics interventions conducted in South Africa between 2010 and 2022. In reviewing these chapters, we are most interested in determining whether we now have a knowledge base to make evidence-based recommendations for how to improve reading and mathematics outcomes in South Africa. We do not automatically assume that interventions are the answer to how one might improve education outcomes in the country. It may well be that improvements are determined by the way that schools are funded, how teachers are allocated to schools, or even broader macroeconomic trends like economic growth or declining unemployment. Yet, despite the importance of these 'non-intervention' factors, we would argue that a synoptic understanding of educational interventions is necessary to understand what does and does not shift learner outcomes. It is also necessary to make evidence-based recommendations about wider reforms. Documenting the contours of the now-substantial evidence base on interventions, its features, regularities, and 'non-negotiables', may yet reveal a productive path for the coming decade.

2 An overview of South African interventions 2010-2022

The period 2010 to 2022 has seen a large growth in the quality and type of both interventions and their evaluations. Perhaps atypically, the South African lineage of interventions began with a large-scale, heavy-dose, province-wide, and government-funded intervention that was then followed by a series of smaller-scale, donor-led interventions testing much the same intervention approach (instructional coaching). The key difference, apart from the scale and source of funding was that the second generation of studies was rigorously evaluated to determine the causal impact on learner outcomes, or a lack thereof. The initial large-scale intervention was the Gauteng Primary Language and Mathematics Strategy (GPLMS) which employed over 500 coaches at a total annual cost to the provincial government of R298-million in 2022 rands (Creecy 2012). A retrospective evaluation using a difference-in-difference methodology found that the intervention led to large gains in both reading and mathematics (see Fleisch & Schoër 2014; Fleisch et al. 2016).

Although this programme was deemed a success by both the evaluation and the government at the time, it did not survive the political transition from one provincial minister of education to another, with the latter shifting in focus from literacy and numeracy to the roll-out of technology. Notwithstanding the changing political winds, the 'triple-cocktail' approach employed in the GPLMS study (lesson plans, materials, and instructional coaches) still had strong local and international proponents (Knight 2007) and was subsequently re-introduced in a different province (the North West) as a medium-scale research project in the form of a partnership between the Department of Basic Education and researchers at the University of Witwatersrand; the Early Grade Reading Study (EGRS). It aimed to determine which intervention approach (coaches, centralised training, or home support) led to the greatest improvement in Setswana Home Language reading in Grades 1-3. Fleisch & Alsofrom in the present volume detail both the intellectual logic of the intervention as well as the lessons learned from its various incarnations and evaluations. Where GPLMS ran from 2010 to 2013, the first EGRS study ran from 2015 to 2018 and the second EGRS from 2017 to 2019. The latter was implemented in Mpumalanga and aimed to improve EFAL reading outcomes. Collectively these studies showed that the only intervention arms that consistently raised learning outcomes were the provision of materials and lesson plans and, crucially, an in-classroom instructional coach. The ratio of coaches to schools was usually 1 coach to 33 teachers. Although the same package of materials combined with centralised teacher training had a small positive impact, other intervention arms (digital coaches, home support) were ineffective and did not lead to improvements in reading outcomes (Cilliers et al 2019). It should also be noted that the home-language intervention (EGRS I) was more successful than the EFAL intervention (EGRS II).

At around the same time as the EGRS study (2015/16) the National Education Collaboration Trust (NECT) was implementing various programmes aimed at improving learning outcomes at scale. These were usually implemented in a large number of schools and within existing resource constraints. For example, in the Programme to Improve Learning Outcomes (PILO), also referred to as the Jika iMfundo programme, "The unit of scale for the pilot was all 1,200 public primary and secondary schools in the two districts" (Metcalfe & Witten 2019, 337), elsewhere referred to as a 'trial at scale' (2019, 340). By contrast the EGRS coaching intervention targeted only 50 schools. The NECT did not include ex-ante plans to evaluate its various programmes, but retrospective reviews of the interventions suggest some improvements in compliance and learner-workbook coverage between intervention districts and non-intervention districts. However, given the non-random assignment of districts, causal claims are not possible. As the independent evaluator concluded in their 2019 report:

> *The results must be read with caution in terms of a causal analysis and attribution to the programme, since there was no pre-intervention baseline and district/school allocation to intervention and control could not be randomised...Results should be interpreted as the difference between treatment and comparison schools, rather than as causal programme impact*
>
> (DNA Economics 2019, 13).

For the purposes of the present chapter the salient features of the NECT and PILO interventions are that they were large-scale (1,000+ schools) and worked within the financial and logistical constraints of the system as it is now, and that they lacked causal evaluations assessing learning outcomes, making it difficult to determine if there were any improvements in learner outcomes.

Another important seam of intervention research that began in the second half of the decade was the Funda Wande series of interventions[4] in the Eastern Cape and Limpopo, both evaluated through randomised control trials with 30-40 intervention schools per arm. The Eastern Cape coaching intervention tested a triple-cocktail model very similar to that of the EGRS I and EGRS II programmes, with similar effect sizes, at least pre-pandemic. As Ardington & Meiring (2020, 66) conclude in their evaluation:

> *To date, the state of evidence from the EGRS I (Cilliers et al. 2019a, 2019b), EGRS II (Kotze et al. 2019) and Funda Wande coaching interventions suggests that structured pedagogy programmes can be effective in more than one province (the North West, Mpumalanga and the Eastern Cape) and in more than one language of instruction (Setswana Home Language, EFAL and isiXhosa Home Language). This lends further external validity to earlier experimental evidence that on-site coaching interventions can shift instructional practice and improve early learning outcomes in low-resource South African schools.*

That effect sizes were so similar when implemented in a different province by a different organisation suggests that these effect sizes (~0.2 SD) are what can be expected from coaching interventions, at least when implemented at a similar dosage by coaches and organisations similar to Class Act (EGRS) and Funda Wande.

In addition to the Eastern Cape coaching intervention, Funda Wande also implemented randomised control trials in Limpopo, testing the impact of a Grade 1 teacher assistant (TA) on both reading and mathematics outcomes during the pandemic (2021). This intervention is important for at least two reasons: (1) it is the first rigorously-evaluated intervention to test an intensive model of support other than coaching – in this case one permanent full-time TA per teacher. And (2) it is the first randomised-control trial evaluating a mathematics intervention in South Africa. Two chapters in the present volume review these two components: Makaluza & Mpeta on the teacher assistant intervention, and Sapire et al. on the mathematics intervention (referred to as Bala Wande). Both show promising results. The impact of the TA intervention on EGRA scores was 0.44 standard deviations after one year (three terms), twice as large as the impact of either the EGRS or FW coaching interventions after one or two years (0.14-0.23 SD) (Ardington & Henry, 2021).

The current volume takes a broad view of interventions and while no book can encompass all interventions across an entire decade, the programmes included in this volume represent the lion's share of both intervention 'types' as well as of those that have shaped the educational zeitgeist in the country. Broadly speaking there are three themes animating the volume. The first deals with the development and implementation of text-based interventions to improve reading outcomes. McKay and Spaull document the genesis, development, and roll-out of the DBE Workbooks. Introduced in 2011 these government-funded books have become a 'minimum floor' resource and are the only ubiquitous resource available in all no-fee schools in the country. They further use the School Monitoring Survey to show that two thirds of Grade 3 teachers in the country (68%) say they use the Workbooks for at least two thirds of their class time,

4 In the spirit of full disclosure it should be noted that one of the authors (Nic Spaull) set up the Funda Wande interventions in both the Eastern Cape and Limpopo.

suggesting that these resources have become the de facto curriculum tool in no-fee schools. Katz and Rees document the development of the Molteno Vula Bula Graded Reader series – a set of phonics-based decodable basal readers now available in all eight African languages. In a companion chapter, Ardington and Spaull use a difference-in-difference methodology to evaluate the provincial roll-out of the Vula Bula Graded Reader Anthologies in the Eastern Cape in 2019 and 2020, showing a statistically significant gain in oral reading fluency for grades affected by the intervention.

The second theme of research is intervention designers and implementers reflecting on their respective interventions and discussing their salient features – how they were designed and implemented as well as their dosage, key change mechanisms, and impact (where evaluations are available). This includes a review of the Early Grade Reading Studies (EGRS I and 2) by Fleisch & Alsofrom reflecting on the feasibility, cost-effectiveness, and mechanisms of on-site instructional coaching. Secondly, Makaluza and Mpeta document the lessons learned from the Funda Wande Limpopo Teacher Assistant intervention, highlighting the importance of rigorous youth recruitment and selection as well as alignment between training, materials, monitoring, and support. Thomsen and her colleagues from Room to Read document their work over the previous decade focusing on developing the skill and habit of reading. The last of the mid-scale reading interventions discussed in the volume is that of Olivier et al. who discuss the lessons learned from a technology-based intervention (Reading Eggs) addressing English as a First Additional Language. Given the increasing appetite for technology-based interventions, it is crucial to learn from those actually implementing these interventions in low-resource settings, rather than accepting aspirational government plans at face value.

The third theme of research is around early grade mathematics interventions addressing topics such as formative assessment (Kanjee & Bhana) and mental mathematics (Venkat & Roberts), or those that address the entire curriculum at a particular grade; for example R-Maths at Grade-R level (Spencer-Smith et al.) or the Bala Wande programme at Grade 1 level (Sapire et al.). Two chapters reflect on the NumberSense materials and interventions, the first by Brombacher and Roberts and the second by Moloi et al. Collectively, these mathematics interventions show the importance of clear alignment between training and materials. Of special interest is that all programmes discussed here introduced a new set of materials to the classroom, requiring teachers to align their teaching to the new materials (whether R-Maths or NumberSense or Bala Wande). We will return to this point later but it is worth noting that all interventions in South Africa that are either believed to be promising, or shown to be so (through causal evaluations), introduce significant additional learner-level resources to the classroom.

Finally, the chapter by Chetty et al. on the Primary School Reading Improvement Programme (PSRIP) of the National Education Collaboration Trust stands apart from the other training-based interventions by virtue of its sheer size and scale. Using a 'train-the-trainer' or cascade model where trainers train subject advisors who then in turn train teachers, the aim is to improve EFAL "within the limits of what is doable at scale" (85). Specifically, the programme targeted 7,000 schools and aimed to work within the budgetary realities of government. Although changes in learner outcomes cannot be deduced from the evaluation (it was not designed to measure causal changes in learner outcomes), lessons learned from working at scale are critical for any programme that aims to move beyond 100 or 200 schools.

For example, Chetty et al. (current volume) document the content-knowledge levels of both Foundation Phase teachers and subject advisors showing pervasively low levels of knowledge that are somewhat impervious to training. Their interventions showed an average improvement of around ten percentage points, shifting average scores from around 40% to 50% (Phase 2) or 30% to 40% (Phase 3). While the change in scores is encouraging, the initial levels of knowledge on basic tests are of particular concern, as are the post-improvement scores – all of which still reflect severe gaps in teacher knowledge. Furthermore, tests reveal that even the government officials meant to train teachers (subject advisors) do not understand the content they are meant to train on. This is a key weakness in the train-the-trainer model using subject advisors. We will turn to this point in the next section.

3 Developing a typology of primary-school interventions in South Africa: Dosage, scalability and impact

Reviewing the different types of interventions in South Africa, the real distinction is less between different types of medium-scale randomised control trials and more between large-scale government-led programmes on the one hand, and small-scale NGO-led programmes on the other. That is not to say that all small-scale programmes are essentially the same since there are non-trivial differences in dosage, type of support, materials provided, and type of evaluation. Yet the discussions and concerns between NGO-providers working in 40-60 schools are similar in a way that discussions between small- and large-scale interventions are not. Many interventions that are well-evaluated and have evidence of impact would be difficult to scale. Similarly, many interventions that are already working at scale lack evidence pointing to improvements in learning outcomes.

To elucidate the differences between interventions we have proposed a rubric (Table 1) including the scale of the programme, type of evaluation, evidence of impact, programme costs, level of support (LTSM and teacher support) and level of government buy-in. Each category is rated from 0 (Non-existent) to 4 (High) across these dimensions with the criteria for the score included in the rubric.

As an illustration of the typology, Table 2 below reports the scores for a series of 14 interventions aimed at improving reading outcomes discussed in various chapters in this volume (there have been more reading interventions than mathematics interventions in South Africa and thus we selected the larger group of interventions). Given the importance of causal evidence of programme impact – i.e. did it improve learning outcomes or not – we double the score for causal impact on learning outcomes. Scores are summed to reveal a total programme score, also reported in the table.

We further summarise these outcomes visually in Figure 1 where the total score (*x*-axis) is plotted against a measure of how many schools the programme was implemented in (*y*-axis). The bubbles are weighted by evidence of impact and the shading of the bubbles indicates the type of evaluation; either where there is no credible control group (grey), a quasi-experimental evaluation (light) or a causal evaluation with a legitimate control group (dark).

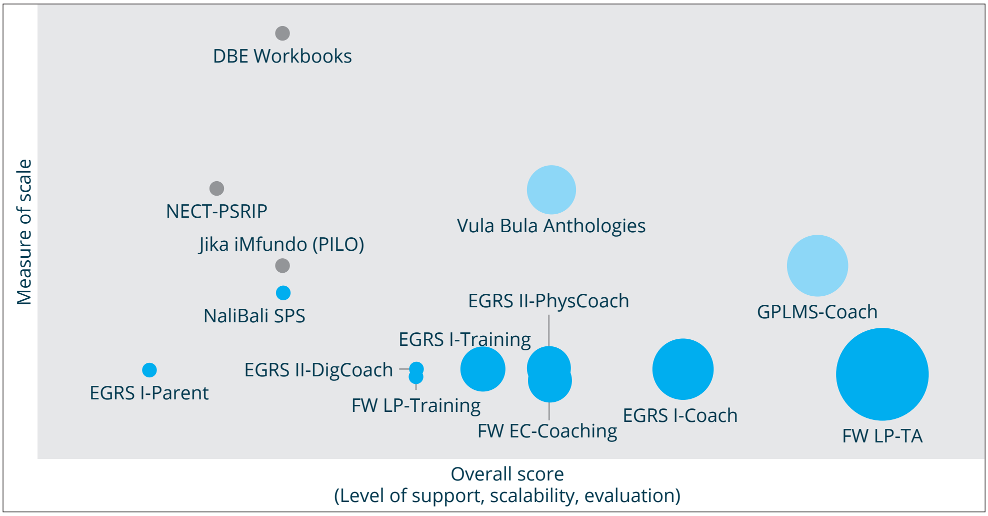
Table 1 Rubric for assessing early grade reading interventions in South Africa

	Element	Rating			
		0 Non-existent	**1 Low**	**2 Medium**	**3 High**
Scale	Location & number of intervention schools	Micro intervention <20 intervention schools	Small-scale intervention (20-200 intervention schools)	Mid-scale intervention (200-1,000 intervention schools)	Large-scale intervention (>1,000 intervention schools)
Additional support	Intensity of centralised teacher training and/or in-classroom support	No centralised training	Any form of centralised training that is less than one full day at least once per term (i.e. four days per year)	Centralised training at least one full day at least once per term by verified expert or in-classroom support at least three times per year by verified expert	In-classroom support or coaching by verified expert at least once per month or more
Additional support	LTSM support Level of materials provision	No additional resources are provided to the school	Additional materials are at the school, teacher or classroom level not the learner level	Additional resources are at the learner level but are not 1-to1 or are not comprehensive	Additional resources are at the learner level, are 1-to-1 & comprehensive (learners are given their own workbooks, readers, manipulatives etc.)
Scalability	Cost per learner: what are the costs of this intervention if it were scaled?	Extremely expensive (> 300% of existing per learner budget)	200-300% of existing per-learner budget	100-200% of existing per-learner budget	Within existing per-learner budgets (i.e. <100%)
Scalability	Level of government ownership / buy-in and funding	Done independently of government	Wholly donor funded but government acknowledgement	Partly government funded (>20% costs paid for by government)	Entirely government led and government funded
Evaluation	Quantitative rigor of evaluation on improvement in learner outcomes	No publicly available evaluation or no measurement of learning outcomes. Data collected by organisation's own employees	Weak. Publicly available evaluation report on learning outcomes but no valid control group. Evaluator selected and contracted by the implementing organisation. Evaluator unknown within the field	Acceptable. Quasi-experimental evaluation. Evaluator selected and contracted by the implementing organisation. Evaluator has strong track record of rigorous evaluations	Ideal. Causal evaluation with valid control group. Evaluator independently selected and contracted by someone other than the implementing organisation. Evaluator has strong track record of rigorous evaluations
Evaluation	Causal evidence of impact of intervention from independent evaluation after up to 2 years of intervention	No impact or no control group	Positive but small impact (<25% of a year of learning)	Moderate-impact (25-50% of a year of learning)	Large impact (>50% of year of learning)

Table 2 Scoring South African early grade reading interventions using the rubric

Intervention	# of intervention schools	Teacher support	LTSM support	Cost	Government buy-in	Evaluation (quantitative rigor)	Causal impact on learning outcomes (rating × 2)	Total score	Reference
EGRS I-Parent	50	0	0	2	1	3	0	6	Cilliers et al. (2020)
NECT-PSRIP	4,198	1	1	3	2	0	0	7	DNA Economics (2019)
Jika iMfundo (PILO)	1,200	1	2	3	2	0	0	8	Metcalfe & Witten (2019)
Nal'ibali SPS	720	1	1	2	1	3	0	8	Ardington et al. (2020)
DBE Workbooks	24,000	0	2	3	3	0	0	8	McKay & Spaull (this volume)
EGRS II-DigCoach	50	1	2	3	1	3	0	10	Cilliers et al. (2022)
FW LP-Training	40	2	3	1	1	3	0	10	Ardington & Henry (2021)
EGRS I-Training	50	1	2	2	1	3	2	11	Cilliers et al. (2020)
FW EC_Coaching	30	2	2	1	1	3	3	12	Ardington & Meiring (2020)
EGRS II-PhysCoach	50	2	2	1	2	3	2	12	Cilliers et al. (2022)
Vula Bula Anthologies	4,298	2	2	3	3	2	2	12	Ardington & Spaull (this volume)
EGRS1-Coach	50	2	2	1	2	3	4	14	Cilliers et al. (2020)
GPLMS-Coach	1,000	2	2	1	3	2	5	16	Fleisch et al. (2016)
FW LP-TA	40	3	3	1	1	3	6	17	Ardington & Henry (2021)

Figure 1 The relationship between scale, support, and impact in South African early grade interventions



Note: *The overall score (x-axis) is the sum of the six components reported in Table 1. The measure of scale is ordinal and categorical with ranges of 1= (20-100); 2=(100-500); 3=(500-2,500); 4=(2,500-5,000); 5=(5,000-10,000), 6=(24,000 all schools)*

There are several important trends evident in Figure 1. Firstly, there seem to be three distinct families of interventions. The first is that of large-scale national interventions that were implemented across provinces and at scale but lack causal evaluations of impact (grey bubbles). Programmes such as the roll out of the DBE Workbooks or the implementation of the NECT's PSRIP programme lack evaluations that can determine whether the programme led to an improvement in learner outcomes as a result of the intervention. The second category (light blue) are province-level interventions (GPLMS in Gauteng, and the Vula Bula Anthologies in the Eastern Cape) which have quasi-experimental evidence. Both evaluations (Fleisch et al, 2016; Ardington & Spaull, this volume) use difference-in-difference methodologies to compare cohorts and grades that received the intervention compared to those that did not. The final group (darker blue) are small scale NGO-led interventions, some of which show causal evidence of impact (bubble size), but all of which operate in 50 schools or fewer and use additional resources not available within the system as it is currently designed, for example instructional coaches or teacher assistants. Yet these are the only studies for which there is bona fide causal evidence of impact on learning outcomes.

3.1 Large-scale trickle-down vs small-scale bottom-up

When reviewing the types of interventions in South Africa documented above, we would argue that two broad categories of intervention emerge. These are (1) a 'train-the-trainer' intervention, and (2) a direct in-classroom support intervention. The former is usually large scale (1,000+ schools) and uses a cascade or 'trickle-down' approach with a very low cost-per-learner. As a result, it is necessarily light-touch. The second approach is usually small-scale (~50 schools), intensive, has a much higher cost-per-learner and involves direct support to teachers in their classrooms. The support could take the form, for example, of an instructional coach or a teaching assistant. Table 3 below summarises and compares some of the key features between the two approaches.

Table 3 Comparing two typical types of education intervention in South Africa 2010-2022

	Trickle-down, train-the-trainer • *Light touch* • *Large scale*	**Bottom-up, in-classroom support** • *Heavy touch* • *Small scale*
Scale	• Large scale with typically more than 1000 schools	• Small scale with typically 30-50 schools in an intervention arm
Exemplar programs	• NECT – PSRIP FP EFAL ~4200 schools; Jika iMfundo (PILO) Whole District Model ~1200 schools	• EGRS I-NW-coaching ~50 schools; FW-EC-Coaching ~30 schools; FW-LP-TAs ~40 schools
Programme leadership & management	• Conceptualised by internal (DBE) or external experts (Class Act), implemented by subject advisors / HODs / existing staff	• Conceptualised by external experts, implemented by external organisation (Class Act, FW, Molteno, R2R).
Implementation	• Train-the-trainer: e.g. each PILO coach is responsible for 110 schools (i.e. >1000 teachers). • Usually uses subject advisors and/or Lead Teachers /HODs with all training being centralised training (e.g. Experts train subject advisors centrally who then train Lead Teachers/ HODs centrally, who then train teachers centrally).	• In-classroom direct support. 1 coach responsible for 33 teachers (EGRS I - NW) or 1 TA per 1 teacher (FW-LP). • Coaches or TAs are externally hired, managed, paid, and monitored by the NGO.
Cost & LTSM	• Low or no-cost model; works largely within existing budgets & uses existing personnel. • Usually teacher-level (curriculum trackers or lesson plans) • If there are learner-level resources, then they are usually low-cost.	• High-cost model; brings in new external support in form of coaches or TAs and procures significant additional LTSM. • Both teacher-level and 1-on-1 learner- level. Teacher-level includes significant Big Books, friezes, classroom library. • 1-to-1 learner-level resources: one anthology per child, one full-color workbook per child, learner-level manipulatives etc.
Monitoring	• Trainings attended and LTSM delivered, e.g.: 5.7 million pieces of teaching and learning materials had been distributed; 84,128 teachers and 671 subject advisors had been trained. • Compliance (curriculum coverage).	• Objectively verifiable standardised assessments (e.g. literacy and numeracy tests for TAs in FW-LP), pre- and post-tests at training (EGRSI-NW), and teacher certification (e.g. Rhodes Adv Cert. for HODs in FW-EC).

	Trickle-down, train-the-trainer • *Light touch* • *Large scale*	Bottom-up, in-classroom support • *Heavy touch* • *Small scale*
Evaluation	• Usually non-existent or ex-post, non-causal, subjective, usually uses no-cost administrative data, no bona fide control schools, no random selection, non-standard instruments, teacher interviews.	• Pre-planned ex-ante, causal, external, independent, learner-level assessments with bona fide control schools and defensible estimates of causal impacts. High cost due to full-service evaluation.
Documentation	• Minimal public documentation. Little publicly available research or evaluation reports.	• Donor-funded evaluations usually publicly available.

All current large-scale programmes in South Africa, primarily NECT and PILO programmes, see subject advisors as key change agents in their theories of change (Metcalfe & Witten, 2019; Chetty et al, current volume). This is primarily because subject advisors are already within the system since they form part of the post-establishment at the province level. Their salaries are covered by existing government resources, and their existing job descriptions include teacher training. Yet there are several reasons why this model of support does not work as planned. We now turn to this point and document it below.

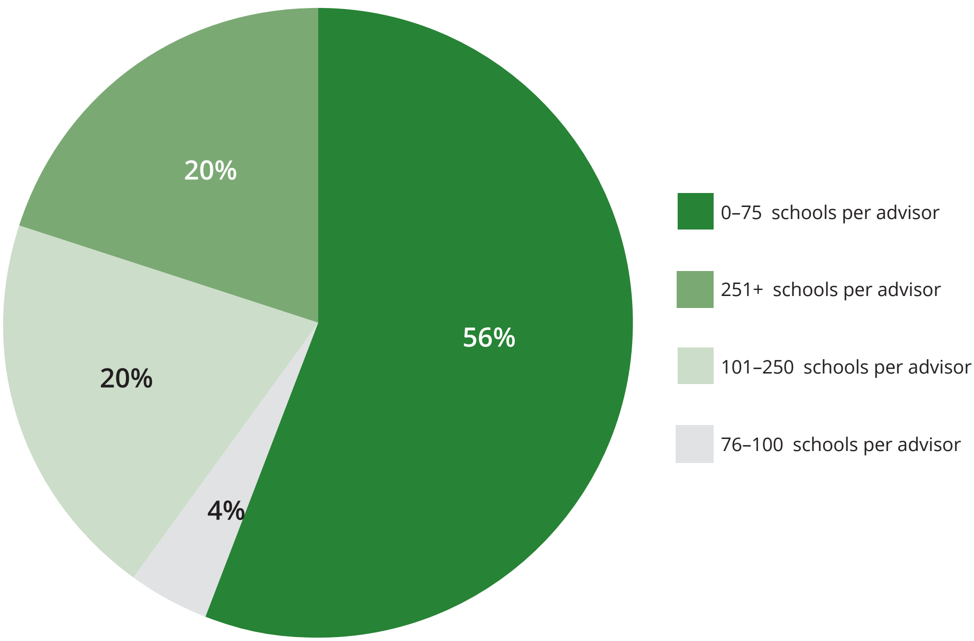
3.2 The number of subject advisors (SAs) and ratio of SAs to teachers

A number of researchers have raised and briefly discussed the high number of teachers or schools to each subject advisor, and that school-based subject-advisor support is consequently infrequent and unrealistic (Taylor 2017; Christie & Monyokolo 2018). In the current volume Chetty et al. also make this point, but explain that their 'train the trainer' model nevertheless employed this approach due to cost constraints and the need to work within the system:

> *While the programme materials are designed in line with term-by-term implementation of CAPS, four training sessions a year (face to face) are simply not a practical, affordable, or efficient use of teachers' time. There is considerable debate about the decision to rely on the existing corps of subject advisors and Foundation Phase specialists who are overstretched and under-resourced in most parts of the system. It is not uncommon for schools in South Africa to say they have not seen a subject advisor for the better part of a year. In that case, why rely on them when school-level professional support to teachers is a known factor in improving outcomes? PSRIP had to make reading improvement workable without the benefit of school level coaching – not by choice, but simply because it was not feasible in the existing institutional or financial parameters of the system*
>
> (Chetty et al. current volume, 73).

However the true extent of how thinly spread subject advisors are remains under-appreciated and somewhat undocumented. A recent NECT-commissioned 'Subject Advisor Profiling Study' sheds new light on this area by surveying all South African subject advisors at the General Education and Training (GET, Gr R-9) Phase for Mathematics and EFAL. Conducted by the Centre for the Advancement of Science and Mathematics (CASME) and published in December 2020, it surveyed all 75 districts in South Africa and reports post establishments, qualifications, and ratios of subject advisors to schools. Figure 2 below illustrates the total number of schools per Foundation Phase subject advisor (Maths and EFAL) for South Africa as a whole (see CASME 2020, 68).

Figure 2 **Total number of schools per Foundation Phase subject advisor (Maths and EFAL)**



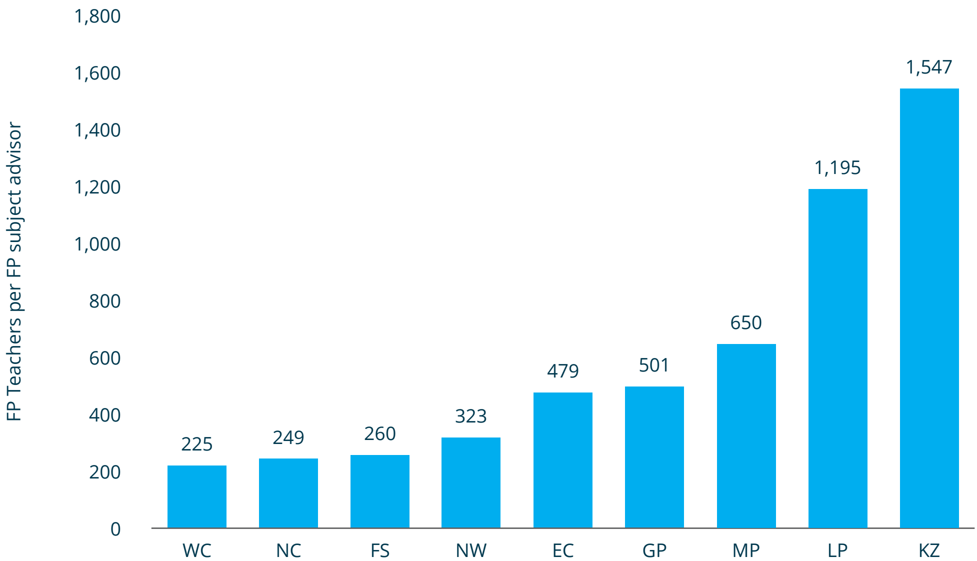
Source: Figure created by N. Spaull & S. Taylor based on information from CASME (2020, 68) report.

> *What is immediately clear is that at least 40% of Foundation Phase EFAL and Mathematics subject advisors are responsible for more than 100 schools making support difficult. Furthermore, subject advisors are expected to provide support to individual teachers within the school. The 2017 Education Labour Relations Council (ELRC) collective agreement on the role of Senior Education Specialists (subject advisors) makes this clear and includes six key performance indicators. Two of these include that they should "conduct regular on-site support visits to teachers in schools" and "Facilitate workshops and training sessions on behalf of their sections/ area of responsibility"* (ELRC 2017, 10).

Given that subject advisors are expected to train and support teachers within the school, the correct ratio to look at is not only subject advisors to schools but subject advisors to teachers. By combining the CASME (2020) data on the provincial post establishments of Foundation Phase subject advisors with two other sources of South African data, on Foundation Phase (1) enrolments and (2) Learner-Educator ratios, this helps establish the true extent of this problem.

Figure 3 below does this by combining the provincial data on Foundation Phase subject advisors[5] from the CASME (2020) study with provincial data on Foundation Phase enrolments (DBE School Realities 2019), as well as provincial Foundation Phase Learner-Educator ratios (Van der Berg et al. 2020, 35).

Figure 3 Average number of Foundation Phase teachers per subject advisor (EFAL and Mathematics) per province



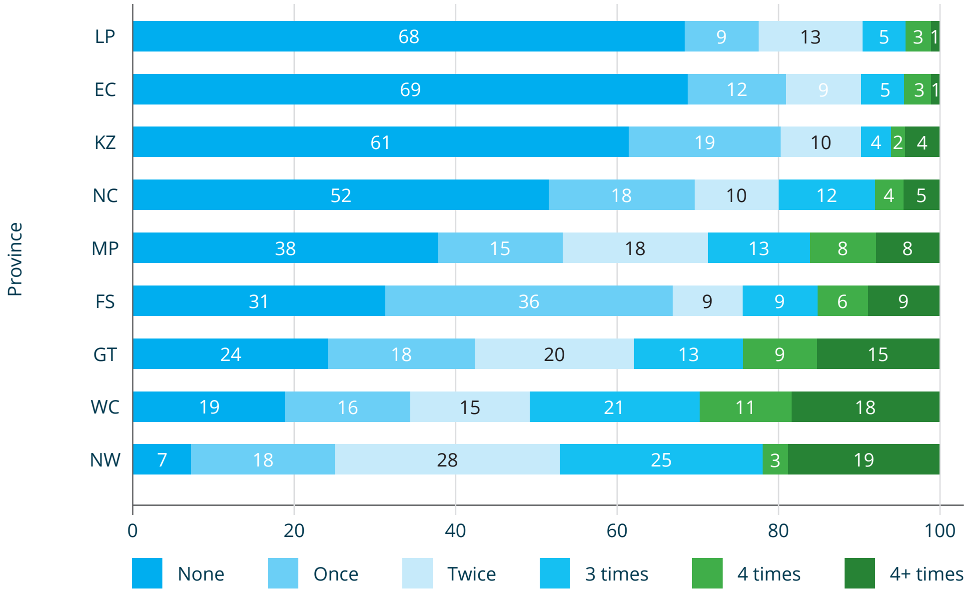
Source: Author's own calculations using information from CASME (2020), DBE (2019), and Van der Berg et al. (2022)

The large inter-provincial variation in teachers per subject advisor reiterates patterns of geographical and resource inequality. In KwaZulu-Natal and Limpopo, the average Foundation Phase subject advisor is responsible for more than 1000 teachers while those in the Eastern Cape, Gauteng, and Mpumalanga for approximately 500 teachers or more.

This survey targeted subject advisors directly and is therefore self-reported data. By contrast the School Monitoring Survey (conducted in 2017) targeted teachers in schools and included questions on when last they had been visited by a subject advisor. Figure 4 reports responses from more than 1543 Grade 3 teachers on the frequency of subject advisor visits in the last year.

5 While this is only for EFAL and mathematics subject advisors, the focal areas of CASME (2020), we have no reason to suspect this would be different for home-language subject advisors.

Figure 4 Percentage of Grade 3 teachers visited by a subject advisor in the previous year.



Source: Own calculations using School Monitoring Survey (2017) data.

The above teacher-reported responses agree with the earlier data but also account for difficulties in visiting schools that are either very small (Eastern Cape) or scattered across a large province (Northern Cape). Two-thirds of Grade-3 teachers in Limpopo, the Eastern Cape and KwaZulu-Natal were never visited by a subject advisor in the previous year. Although frequency of subject advisor visits is higher in the other six provinces, still only about one-in-five teachers were visited at least once per term (four times or more in the year), considered to be the 'bare minimum' in heavy-touch interventions like EGRS or FW.

In addition to the above concerns around the number of subject advisors, there are legitimate concerns around the quality of the support that they can offer, even when they do manage to visit schools. Of the 308 GET Maths and EFAL subject advisors who reported their academic qualifications in the Subject Advisor Profiling Study, 90% had less than a Master's degree and 37% had less than an honours degree (CASME 2020, 48). It should also be noted that basic qualifications are often a necessary (but not sufficient) condition for content knowledge or teacher quality; while most teachers in South Africa meet minimum qualification criteria, they still lack primary-school-level content knowledge in mathematics (Venkat & Spaull 2015). Furthermore, as reported above, when subject advisors are tested on basic content knowledge and curriculum requirement tests, they are often not more knowledgeable than the teachers they are training, frequently scoring 50% on content they are required to train on (Chetty et al. current volume, p. 75).

4 The 'low ceiling' hypothesis

The discussion in the previous section shows quite clearly that there is an implicit trade-off between scale and impact, a point that is well-documented in the literature: "One of the most consistent findings in the education literature is that effects decrease when smaller targeted programs are taken to scale" (Slavin & Smith 2009, in Kraft 2020, 6).

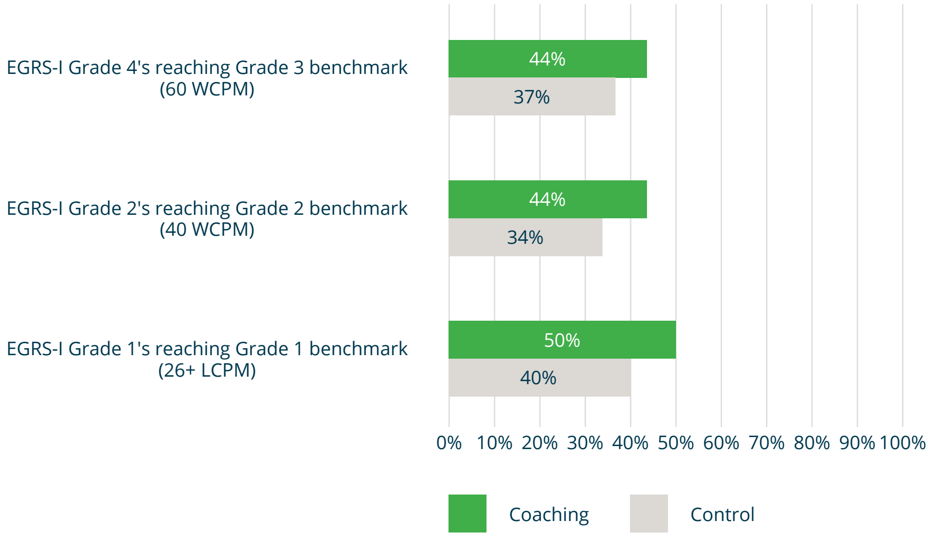
In South Africa, as elsewhere, large-scale interventions are usually either not evaluated or their evaluations are not designed to measure causal changes in learning outcomes. However, even the most successful programmes, and ones that are designed to measure causal impacts in learner outcomes – EGRS and FW – often yield improvements that are more modest than many education stakeholders would anticipate given the intensity of the interventions. For example, both EGRS I (North West) and Funda Wande (Eastern Cape) interventions employed teacher coaches in a ratio of approximately one coach to five to ten schools depending on how many grades the intervention was being implemented in (typically one coach to 30 teachers). As mentioned previously, both interventions yielded improvements of approximately 0.2 standard deviations, an effect size that Kraft refers to as 'large' (2020, 7).

Understanding what these effect sizes mean in practical terms is important if one is to communicate these findings to policymakers, but also to anchor expectations on empirical realities of what is possible, or at least what has been historically possible to date in South Africa. Figure 5 below uses the new language-specific benchmarks for oral reading fluency (Mohohlwane et al. 2022) and letter-sound knowledge[6] to illustrate programme impacts from Funda Wande[7] at the Grade 1 level (Ardington & Meiring 2020) and EGRS I at the Grade-2 and Grade-4 level (Cilliers et al. 2019). Across all three grades the teacher-coaching interventions led to statistically significant gains in the percentage of learners reaching the grade-specific benchmarks of approximately ten percentage points (40% to 50% for the Grade-1 benchmark, 34% to 44% for the Grade-2 benchmark, and 37% to 44% for the Grade-3 benchmark). While it is encouraging that these programmes are sufficiently intensive to lead to measurable causal changes in outcomes, even after three years of a coaching intervention less than half of Grade-4 learners in these schools (44%) could reach the Grade-3 benchmark, only marginally higher than learners in the control school (37%).

6 Here we use the 26+ Letters Correct Per Minute (LCPM) 'Letter Decoder' threshold as in Pretorius & Spaull (2022) rather than the higher 40 LCPM 'Letter Mastery' benchmark as in Mohohlwane et al. (2022).

7 The Ardington & Meiring (2020) evaluation was after only one year of intervention before the Covid-19 pandemic. Subsequent evaluations (Ardington & Henry, 2021) reveal that many of the pre-Covid-19 gains were eroded during the pandemic. The EGRS I intervention was implemented from 2015-2018 and thus not affected by the pandemic.

Figure 5 **Percentage of Grade 1, 2, and Grade 4 learners reaching oral reading fluency benchmarks by treatment arm (EGRS I and Funda Wande coaching interventions)**



Arguably the most intensive intervention implemented in South Africa to date has been the Funda Wande Limpopo Teacher Assistant intervention (Makaluza & Mpeta, current volume) which allocated a full-time paid teacher assistant to each Grade 1 teacher in intervention schools as well as additional materials and termly training. The impact evaluation after one year of intervention showed overall effect sizes twice as high (0,42 SD) as those of the EGRS and FW-EC coaching interventions (Ardington & Henry, 2021). In the 'business-as-usual' control schools in the Limpopo intervention, 80% of Grade 1 learners could not read a single word in Term 4 of Grade 1. Yet even in the 'high intensity high impact' teacher-assistant-intervention arm, 67% of Grade 1 learners could not read a single word after one year of the intervention (Ardington & Henry 2021, 4). Again, it is difficult to celebrate this result when two thirds of Grade 1 learners are still entirely illiterate after one year of schooling, albeit that this intervention was during Covid-19 school closures and rotational timetables in 2021.

Given these results it seems reasonable to ask why these relatively intensive interventions are still yielding only modest improvements in outcomes, even before the complications of the pandemic. One possible explanation referred to earlier in the chapter is that of a 'low ceiling' hypothesis where the overarching superstructures of low teacher content knowledge, high unemployment, and weak accountability relationships persist unabated. While there are other factors that undoubtedly affect learner performance, such as large class size (Spaull et al. 2022a) or a lack of textual resources, unless the prior superstructures are addressed at the same time, it is likely that no 'radical' improvement in learner outcomes is possible. It may well be that 'capacity-based' interventions that lack any forms of accountability reach a low ceiling that they cannot move beyond.

5 Conclusion

The chapters included in this volume represent more than a decade of interventions and evaluations on early grade reading and mathematics. When seen collectively and synoptically it is clear that scholarship on South African interventions has advanced apace in the last decade. Yet it also generates as many questions as it answers. Why are promising small-scale interventions never scaled up? Why are large-scale tax-funded interventions still not evaluated to determine causal impacts on learner outcomes? Why are 'trickle-down train-the-trainer' interventions still the norm when it is clear to all stakeholders that the underlying assumptions (that subject advisors will train teachers) are fundamentally flawed? The ratio of schools to subject advisors is simply too large for them to fulfill any meaningful support function, at least in the vast majority of provinces. Yet over and above this inadequate headcount, the appointment and promotion of subject advisors is not primarily based on skill or competence. How else does one explain why subject advisors lack basic knowledge of either curriculum policy or how to teach reading – frequently scoring 50% on tests aimed to measure such knowledge (Chetty et al. current volume)?

Trickle-down train-the-trainer models are usually justified due to the cost-constraints faced by national and provincial government – constraints that are real and cannot be wished away. Yet the only interventions for which there is evidence of impact cost more than the current cost-per learner in no-fee schools. This may mean one of two things:

1. Successful small-scale interventions will not be scaled up at government expense until there is a greater willingness to re-allocate budgets, which is dependent on political will and the perceived importance of foundational reading and mathematics. Not since the GPLMS intervention in 2010 has any province implemented a province-wide tax-funded early grade reading or mathematics intervention that could credibly be referred to as substantive or intensive. The 'scale-up' of the EGRS study in the North West (called the Reading Support Program) is still primarily donor-funded, works in only 263 schools and chose to implement a less intensive version of the model with only 140 schools receiving coaching from 14 coaches (Williams et al. 2019).
2. There is a need to invest in a new type of intervention. All interventions covered in this volume, and implemented over the last decade, are 'capacity' interventions aimed at either creating and distributing new materials (graded readers, workbooks, manipulatives) or providing additional support and training to teachers and government officials. In Elmore's Capacity-or-Accountability dichotomy (Spaull 2015) all South African interventions in the last decade, whether small-scale or large, eschew accountability and propose capacity-based solutions. In our view this is largely due to a pragmatic consideration of the current political climate in the country and working 'within the realm of the possible'. Interventions aimed at creating any accountability mechanisms would be vociferously opposed by the politically powerful teacher unions. An example of such an intervention is offering intensive training and in-classroom support to teachers together with teacher evaluations leading to defensible consequences for non-performance. It reflects the South African discourse on teachers that there are no serious policy analysts or researchers who have proposed replacing the least knowledgeable teachers who

cannot score even 30% on a primary-school content-knowledge test, even after repeated training and second-chance opportunities. Yet how can teachers teach if they lack the most elementary knowledge about mathematics or reading?

As we reflect on the chapters in this volume and look ahead to the next decade of intervention research, the scholarly community will need to address the outstanding questions related to capacity interventions discussed in the present volume (scaling small-scale interventions and reformulating large-scale interventions), but also conceptualise and experiment with a new generation of intervention aimed at improving accountability and addressing the superstructure of weak teacher content knowledge and quality. At the end of Grade 1, more than half of learners in no-fee schools do not know the letters of the alphabet (Wills et al. 2022), and they cannot add and subtract single-digit numbers (Spaull et al. 2022a), despite a year of formal full-time schooling. While factors like extreme class sizes undoubtedly influence these outcomes, the figures in 'acceptable' class sizes of 35 learners per teacher are only marginally better. The low quality of teachers and, importantly, the low quality of teaching is arguably one of the leitmotifs running through all three volumes published in this series.

New research on teacher demographics (Van der Berg et al. 2022) reveals that half (45%) of all publicly employed teachers will retire in the next 10 years (National Treasury 2021, 63) creating an unprecedented opportunity to recruit, train, and employ a new generation of teachers. It is rare for such demographic opportunities to emerge, and this is an opportunity that could help address one of the underlying superstructures behind persistently weak outcomes.

References

Ardington, C. and Henry, J. (2021). Impact Evaluation of Funda Wande Teacher Assistant Intervention in Limpopo Province. Southern Africa Labour and Development Research Unit. Cape Town: University of Cape Town.

Ardington, C. and Meiring, T. (2020). Impact Evaluation of Funda Wande Coaching Intervention Midline Findings. Southern Africa Labour and Development Research Unit. Cape Town: University of Cape Town.

Ardington, C., Hoadley, U., & Menendez, A. (2019). Impact Evaluation of USAID/South Africa Story Powered School Program Endline Report. Washington, DC: USAID Reading and Access Activity.

Ardington, C., Wills, G., & Kotze, J. (2021). COVID-19 learning losses: Early grade reading in South Africa. *International Journal of Educational Development*, 86.

Armstrong, P. (2015). Teacher Characteristics and Student Performance: An Analysis Using Hierarchical Linear Modelling. *South African Journal of Childhood Education*, 5(2), 123–145.

Awich, M. (2021). SACMEQ IV International Report. Southern and Eastern African Consortium for Monitoring Education Quality (SACMEQ). Available at: http://www.sacmeq.org/sites/default/files/sacmeq/reports/sacmeq-iv/international-reports/sacmeq_iv_international_report.pdf (Accessed 29 March 2022).

Bowie, L., Venkat, H., & Askew, M. (2019). Pre-service primary teachers' mathematical content knowledge: An exploratory study. *African Journal of Research in Mathematics, Science and Technology Education*, 23(3), 286–297.

Centre for the Advancement of Science and Mathematics Education (CASME). (2020). Subject Adviser Profiling Study South Africa. Pretoria: National Education Collaboration Trust & Department of Basic Education.

Christie, P., & Monyokolo, M. (Eds.). (2018). Learning about sustainable change in education in South Africa: The Jika iMfundo campaign 2015-2017. Johannesburg: Saide.

Cilliers, J., Fleisch, B., Prinsloo, C., & Taylor, S. (2019). How to improve teaching practice? An experimental comparison of centralized training and in-classroom coaching. *Journal of Human Resources, 55*(3), 926–962.

Cilliers, J., Fleisch, B., Kotze, J., Mohohlwane, N., Taylor, S., & Thulare, T. (2022). Can virtual replace in-person coaching? Experimental evidence on teacher professional development and student learning. *Journal of Development Economics,* 102815.

Creecy, B. (2012). Gauteng Education Budget Vote 2012/13. Available at: https://www.gov.za/gauteng-education-budget-vote-201213 (Accessed 14 November 2022).

DBE. (2016). Report of The Ministerial Task Team Appointed by Minister Angie Motshekga to Investigate Allegations into the Selling of Posts of Educators by Members of Teachers Unions and Departmental Officials in Provincial Education Departments: Final Report.

DBE. (2018). Mathematics Teaching and Learning Framework for South Africa: Teaching Mathematics for Understanding. Pretoria: Department of Basic Education.

DBE (2019). School Realities 2019. Pretoria: Department of Basic Education.

DBE. (2020a). National Framework for the Teaching of Reading in African Languages in the Foundation Phase. Pretoria: Department of Basic Education.

DBE. (2020b). Action Plan to 2030. Department of Basic Education. Available at: https://www.education.gov.za/Portals/0/Documents/Publications/Sector%20plan%202019%2015%20Sep%202020.pdf?ver=2020-09-16-130709-860 (Accessed 14 November 2022).

DNA Economics. (2019). Impact and Process Evaluation of NECT Learning Programme (Foundation Phase). Final Results Technical Report. 8 July 2019.

Education Labour Relations Council. (2017). Job Descriptions for Office Based Educators. (Accessed 15 October 2020) https://nationalgovernment.co.za/entity_annual/1194/2017-education-labour-relations-council-(elrc)-annual-report.pdf (Accessed 14 November 2022).

Fleisch, B. & Schöer, V. (2014). Large-scale instructional reform in the Global South. *South African Journal of Education*. 23 (4).

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

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. University of Pretoria. Pretoria: Centre for Evaluation and Assessment.

JET. (2016). Report on the Implementation Evaluation of the National School Nutrition Programme. Pretoria: Department of Performance Monitoring and Evaluation (DPME).

Knight, J. (2007). *Instructional Coaching: A partnership approach to improving instruction*. Thousand Oaks: Corwin Press

Kraft, M. A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher,* 49(4), 241–253.

Metcalfe, M. and Witten, A. (2019). Taking change to scale: Lessons from the Jika iMfundo campaign for overcoming inequalities in schools. In Spaull, N. & Jansen, J. (eds): *South African Schooling: The Enigma of Inequality*. Cham: Springer.

Mohohlwane, N., Ardington, C., & Wills, G. (2022). A review of recent efforts to benchmark early reading skills in South African languages. In Spaull, N. & Pretorius, E. (Eds) *Early Grade Reading in South Africa*. Cape Town: Oxford University Press.

Mullis, I.V.S., Martin, M.O., Foy, P., & Arora, A. (2012). *TIMSS 2011 International Results in Mathematics*. Chestnut Hill: TIMSS Study Center Boston College.

National Treasury. (2021). Medium Term Budget Policy Statement 2021. Pretoria: National Treasury.

Nuga Deliwe, C., & Van der Berg, S. (2022). The role of assessment in Foundation Phase improvement: The Annual National Assessments and beyond. In Venkat, H. & Roberts, N. (Eds) *Early Grade Mathematics in South Africa*. Cape Town: Oxford University Press.

Pretorius, E., Rastle, K., & Mtsatse, N. (2022). A curriculum review of South African CAPS Grades R-3. In Spaull, N. & Pretorius, E. (Eds) *Early Grade Reading in South Africa*. Cape Town: Oxford University Press.

Pretorius, E., & Spaull, N. (2022). Reading research in South Africa (2010–2022): Coming of age and accounting for empirical regularities. In Spaull, N. & Pretorius, E. (Eds) *Early Grade Reading in South Africa*. Cape Town: Oxford University Press.

Reddy, V., Kanjee, A., Diedricks, G., & Winnaar, L. (2003). Mathematics and Science Achievement at South African Schools in TIMSS 2003. Human Sciences Research Council. Available at: https://www.timss-sa.org/publication/mathematics-and-science-achievement-at-south-african-schools-in-timss-2003 (Accessed 1 March 2022)

Reddy, V., Winnaar, L., Juan, A., Arends, F., Harvey, J., Hannan, S., Namome, C., Sekhejane, P., & Zulu, N. (2020). *TIMSS 2019 Highlights of South African Grade 9 Results in Mathematics and Science*. HSRC.

Roberts, N., Feza, N., & Graven, M. (2022). Story telling in early grade mathematics classrooms in South Africa. In Venkat, H. & Roberts, N. (Eds) *Early Grade Mathematics in South Africa*. Cape Town: Oxford University Press.

Shepherd, D., & Mohohlwane, N. (2021). The impact of COVID-19 in education – more than a year of disruption. National Income Dynamics Study Coronavirus Rapid Mobile Survey (NIDS-CRAM). Cape Town: NIDS-CRAM.

Spaull, N. (2015). Accountability and capacity in South African education. *Education as Change, 19*(3), 113–142.

Spaull, N., Lilenstein, A., & Carel, D. (2020). *The Race Between Teacher Wages and the Budget: The Case of South Africa 2008–2018. Research on Socioeconomic Policy (RESEP)*. Stellenbosch: RESEP

Spaull, N., Courtney, P., & Qvist, J. (2022). Back to the beginning: Growing class sizes and stagnating mathematics learning outcomes in TIMSS South Africa (Grade 5, 2015-2019). In Venkat, H. & Roberts, N. (Eds) *Early Grade Mathematics in South Africa*. Cape Town: Oxford University Press.

Taylor, N. (2017). Report on the Implementation Evaluation of the National Curriculum Statement Grade R to 12 Focusing on the Curriculum and Assessment Policy Statements (CAPS). Pretoria: Technical Report.

Taylor, N. (2021). The dream of Sisyphus: Mathematics education in South Africa. *South African Journal of Childhood Education,* 11(1).

Taylor, N., & Mawoyo, M. (2022). Professionalising teaching: The case of language and literacy. In Spaull, N. & Pretorius, E. (Eds) *Early Grade Reading in South Africa*. Cape Town: Oxford University Press.

Van der Berg, S., Gustafsson, M., & Burger, C. (2022). School Teacher Supply and Demand in South Africa in 2019 and Beyond. Pretoria: Department of Higher Education and Training.

Van der Berg, S., Gustafsson, M., & Burger, C. (2020). School Teacher Supply and Demand in South Africa in 2019 and Beyond. Pretoria: Department of Higher Education and Training

Van der Berg, S., Patel, L., & Bridgman, G. (2022a). Food insecurity in South Africa: Evidence from NIDS-CRAM wave 5. Development Southern Africa.

Van der Berg, S., Hoadley, U., Galant, J., van Wyk, C., & Böhmer, B. (2022b). *Learning Losses from COVID-19 in the Western Cape: Evidence from Systematic Tests*. Research on Socioeconomic Policy (RESEP). Stellenbosch: RESEP

Van Staden, S., & Gustafsson, M. (2022). What a decade of PIRLS results reveals about early grade reading in South Africa: 2006, 2011, 2016. In Spaull, N. & Pretorius, E. (Eds) *Early Grade Reading in South Africa*. Cape Town: Oxford University Press.

Venkat, H., & Roberts, N. (2022). Conclusion: All children calculating with confidence and understanding by 2030: What will it take? In Venkat, H. & Roberts, N. (Eds) *Early Grade Mathematics in South Africa*. Cape Town: Oxford University Press.

Venkat, H., & Spaull, N. (2015). What do we know about primary teachers' mathematical content knowledge in South Africa? An analysis of SACMEQ 2007. *International Journal of Educational Development*, 41, 121–130.

Williams, B., Bisgard, J., Tjasink, K., Taimo, L., Roper, M., Dornbrack, J., Orr, J., & Baba, N. (2019). Reading support project: Final design evaluation report: Design evaluation of the reading support project. USAID.

Wills, G., Ardington, C., & Lesang Sebaeng, M. (2022). Foundational skills in home language reading in South Africa:Empirical evidence from 2015–2021. In Spaull, N. & Pretorius, E. (Eds) *Early Grade Reading in South Africa*. Cape Town: Oxford University Press.

Zondo, R. M. M. (2022). Judicial Commission of Inquiry into State Capture Report.

Zuze, L., Reddy, V., Visser, M., Winnaar, L., & Govender, A. (2018). TIMSS 2015 Grade 9 National Report: Understanding Mathematics and Science Achievement Amongst Grade 9 Learners in South Africa. Available at: TIMSS 2015 Grade 9 National Report – TIMSS SA (Accessed 14 November 2022).

02

Changing the 'grammar of schooling' in South Africa: The case of the DBE Workbooks

VERONICA MCKAY & NIC SPAULL

Abstract

In this chapter we provide an overview of the genesis, development, and roll-out of a South African workbook initiative, the Department of Basic Education (DBE) Workbooks. This is a nation-wide government-funded programme that provides full-colour curriculum-aligned workbooks for core primary-school subjects. (home language, mathematics, life skills, and English first additional language). Over 65-million workbooks are distributed to more than 24,000 schools each year at an annual cost of approximately R1-billion. We illustrate how political continuity and economies of scale have led to the creation of a new 'minimum floor' of text resources in every South African classroom. We further draw on two nationally representative surveys – a household survey and a school survey – both of which attest to their ubiquity across every province. We report on a previously under-analysed survey (School Monitoring Survey of 2017) of more than 1,000 Grade 3 teachers' responses to questions on the DBE Workbooks on their utility in schools. The findings show that 79% of Grade 3 teachers are satisfied with the Workbooks, 84% find them helpful and at least two thirds of teachers (68%) say they use the DBE Workbooks for at least two thirds (66%+) of their class time. These findings suggest that the DBE Workbooks have become the de facto curriculum tool in South African schools and we argue that they are a rare exception of a resource-led change to the 'grammar of schooling' in South Africa. It is likely that they contributed to national gains in learning outcomes over the decade 2010-2019.

KEYWORDS
WORKBOOKS
resource-led learning, Curriculum and Assessment Policy Statement (CAPS), multilingualism, literacy, curriculum reform

1 Introduction

The tale of post-apartheid South African schooling is essentially a story in two parts. The first covers roughly the first two decades of democracy (1994-2011) and is dominated primarily by radical curriculum reform, a de-racialisation and teacher pay equity (Muller & Hoadley 2019; Fiske & Ladd 2014; Gustafsson 2019). The second part, covering nearly ten years (2011-2019) has seen the locus of attention shift away from bureaucratic policy discussions and academic curriculum disagreements towards a growing coherence around four questions related to what actually happens in the classroom: (1) what do teachers teach? (2) how do they teach it? (3) what do they use? and (4) how do they assess learning? Broadly speaking these relate to curriculum, pedagogy, materials, and assessment. Building on Tyack and Tobin (1994) we refer to these as the 'grammar of schooling' and argue that the increased attention paid to these fundamental practical components of schooling – particularly in the last decade – were beginning to yield results in improved learning outcomes, at least before the Covid-pandemic. However, before discussing these improvements and what may lie behind them, it is worth clarifying what we mean by the 'grammar of schooling':

> *The basic 'grammar' of schooling, like the shape of classrooms, has remained remarkably stable over the decades. By the 'grammar' of schooling we mean the regular structures and rules that organize the work of instruction (Continuity in the grammar of instruction has frustrated generations of reformers who have sought to change these standardized organizational forms) we ask why this grammar persisted and why challenges generally did not succeed. Practices like graded classrooms structure schools in a manner analogous to the way grammar organizes meaning in languages. Neither the grammar of schooling not the grammar of speech needs to be consciously understood to operate smoothly. Indeed, much of the grammar of schooling has become so well established that it is typically taken for granted as just the way schools are. It is the departure from customary practice in schooling or speaking that attracts attention*
>
> (Tyack & Tobin, 454).

We argue that while there has been a steady stream of well-intentioned reforms in post-apartheid South Africa, few of these have 'stuck' sufficiently to change the grammar of schooling. The aim of this chapter is to document one we believe has – the introduction of the *DBE Rainbow Workbooks* (hereafter 'the DBE Workbooks').

Despite their ubiquity and for reasons that are not well understood, this decade-long, country-wide intervention has received minimal scholarly attention to date. We aim to remedy this situation by drawing on two primary sources of information: (1) first-hand knowledge of the rationale, development, funding, and implementation of the intervention. We believe we are well-placed to do this since one of the authors of this chapter (Veronica McKay) was the Overall Project Coordinator for the DBE Workbooks and to that effect was seconded to the Department of Basic Education from 2008 to 2012; and (2) analysing a previously under-analysed dataset, the School Monitoring Survey (SMS) of 2017 which is an independently administered and nationally representative sample of 992 South African primary schools which included a specific sub-questionnaire regarding the reported use of the DBE Workbooks by teachers and learners, including teachers' perceptions of the quality, helpfulness and sufficiency of these workbooks.

We begin by providing background information on the DBE Workbooks and their development. Thereafter we review the limited literature pertaining to the workbooks, as well as the grammar of schooling in SA. Thereafter we provide an analysis of the 2017 School Monitoring Survey data in relation to the preceding literature. Finally we conclude the chapter by foregrounding the relevance of this work and discussing the way forward.

2 Background

2.1 Rationale, development, and implementation of the DBE Workbooks

The idea of developing a national series of grade-specific workbooks for literacy and numeracy was relatively widespread in South African policymaking circles in 2008 and 2009. In March 2008, minister of education Naledi Pandor launched the Foundations For Learning (FFL) campaign, "a four-year campaign to create a national focus to improve the reading, writing, and numeracy abilities of all South African children" (South African Parliament 2008, 4) and recommended that "A Workbook or Learners' book (to provide systematic development of literacy, phonic and language skills" be developed (12). But the de facto genesis of the DBE Workbook project can be found in a presidential proclamation two years after the FFL policy was promulgated. In the president's State of the Nation Address (February 2010) President Zuma announced two important education reforms that had long been mooted by the Department of Education: a new workbook series (which would become the DBE Workbooks) and a set of national primary school assessments (which would become the Annual National Assessments, ANAs):

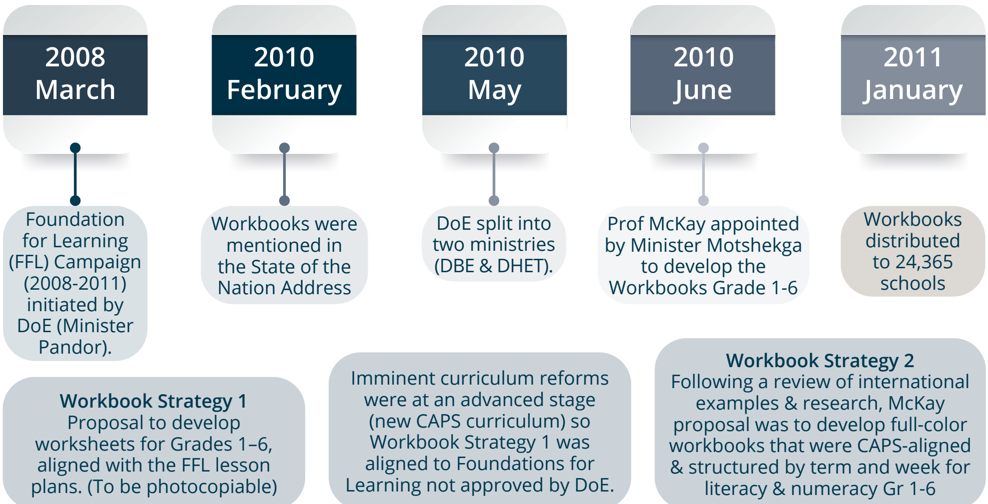
> *We want to improve the ability of our children to read, write and count in the foundation years. Unless we do this, we will not improve the quality of education. Our education targets are simple but critical… We will assist teachers by providing detailed daily lesson plans. To students, we will provide easy-to-use workbooks in all 11 languages. From this year onwards, all grade 3, 6 and 9 students will write literacy and numeracy tests that are independently moderated*
>
> (Zuma, 2010).

Shortly after the 2010 State of the Nation Address, the Department of Education's FFL presented a Workbook Strategy (2010) with the workbooks being perceived as an extension to the FFL which had been launched in 2008 and with the intention of ensuring that by 2011 all learners would be able to demonstrate age-appropriate levels in literacy and numeracy. The FFL had provided assessment frameworks and lesson plans for teaching literacy and numeracy for Grades 1-6. The FFL's 'worksheet strategy' was to develop worksheets that were aligned to the FFL lesson plans (DoE 2010) envisaging 200 photocopiable worksheets for both language and mathematics – referred to as 'Workbook Strategy 1' in the timeline below.

This proposal did not find support with the DoE, due to the process to replace the National Curriculum Statement (NCS) with the new Curriculum and Assessment Policy

Statement (CAPS). Although the CAPS was only formally introduced in schools in 2012, by 2010 the plans were already at an advanced stage and the DoE decided that it would be unwise to publish workbooks aligned to the old FFL materials and the old curriculum, since this would be out of sync with the imminent roll out of the new curriculum[1].

Figure 1 **Timeline of significant dates for the initiation of the DBE Workbooks**



In April 2009, President Zuma announced that the Department of Education would be split into two departments: the Department of Higher Education and Training (DHET) and the Department of Basic Education, addressing higher education and schooling respectively. Minister Angie Motshekga, former MEC for Education in the Gauteng Department of Education, was appointed as the minister of Basic Education in 2009, a position she has subsequently been re-appointed to twice, and one she continues to hold in 2022 making her one of the longest serving ministers in the cabinet.

As part of this reorganisation, the successful Kha Ri Gude Literacy (KRG) Campaign[2] was relocated from the DoE to the newly established DBE. In June 2010, Minister Motshekga requested the CEO of Kha Ri Gude, Prof Veronica McKay, to take on the task of conceptualising and coordinating the development of the school workbooks for languages and mathematics for Grades 1-6. The KRG Campaign had successfully

1 The revisions of the National Curriculum Statement (NCS) for Grades R-12 were formally promulgated on 11 September 2011 in Government Gazette 722 of No. 34600 when the CAPS was approved as the new curriculum for all subjects for Grades R-12. The three reforms of a new curriculum (CAPS), a new national primary school assessment system (ANAs), and a new national set of curriculum materials (DBE Workbooks) were all formally implemented at approximately the same time in 2011/2012.

2 Kha Ri Gude was a mass adult literacy campaign from 2008 to 2017 that reached 4.7 million adults with the aim of teaching basic literacy and numeracy skills equivalent to the third grade of schooling (McKay 2013, 2015, 2020). The development of the campaign's literacy approach was based on substantial research into literacy and language development which was valuable for the development of the workbooks.

developed workbooks[3] in all 11 official languages, based on extensive research into literacy learning, and had an established team of African language experts with considerable experience in materials development. McKay was required to submit a new proposal for workbooks to meet the target of January 2011 implementation (DBE 2010b).

The new strategy based on the earlier research into literacy and language learning and was influenced by international studies where workbooks had been used. While the DBE originally conceptualised photocopiable loose-leaf worksheets, McKay's revised strategy (2010b) took account of the challenges teachers had in copying and distributing resources. Drawing on her extensive Unisa distance education experience and the distribution model of KRG, she developed a proposal for full colour workbooks aligned by term and week with the CAPS, and colour-coded for ease of distribution, hence the 'Rainbow' Workbooks. McKay's submission to the Ministerial Management Committee[4] in 2010 provided mock-up worksheets that were approved. This is referred to as 'Workbook Strategy 2' in the timeline.

2.2 Research informing the DBE Workbooks

Although the time period for the development of the workbooks was short, their development drew upon the literacy and numeracy research previously conducted for KRG as well as on an extensive review of existing large-scale local and international assessments.

2.2.1 The extent of the problem: Reviewing learning outcomes from large-scale assessments

By 2010 South Africa had participated in a number of large-scale assessments of learning outcomes at the primary-school level. This included local assessments such as the Systemic Evaluations (of 2001 and 2007 that tested Grade 3 literacy and numeracy in representative samples of more than 2,000 schools (DoE 2008); and the longitudinal National School Effectiveness Study (NSES) that tested literacy and numeracy in Grades 3, 4, and 5 in 2007, 2008, and 2009 in more than 200 schools (Taylor 2011). These studies found literacy and numeracy scores ranging between 20-35% with minimal gains in learning outcomes across the system over time. The ANA Diagnostic of 2009 confirmed these findings and provided more granular detail on the curriculum areas that learners (and by implication their teachers) found difficult.

The Progress in International Reading Literacy Study (PIRLS, Grade 4 2006) revealed that 87% of Grade 4 learners were at risk of not learning to read (Howie et al. 2008), while the Southern and Eastern African Consortium for Monitoring Educational

3 The literacy campaign received recognition for for its research-based materials: UNESCO Confucius Literacy Award (2016): the National Ubungcweti GCIS (2009) and the PANSALB Award (2010) for materials in 11 languages, the introduction of SA Sign Language, and for materials being produced in Braille.

4 The Ministerial Management Committee (MMM) comprises the minister and deputy minister, ministerial advisors, the director general, and the various deputy director generals of the national Department of Basic Education.

Quality (SACMEQ, Grade 6 2000 and 2007) showed that there had been no progress at the national level in either literacy or numeracy between 2000 and 2007 (Moloi & Chetty 2011, 51; DBE, 2011d).

2.2.2 The South African schooling context

The literature on teaching and learning in South African classrooms offered numerous insights that were significant for the development of the workbooks. The first of these was weak teacher-content knowledge and pedagogical-content knowledge (PCK) (Carnoy & Chisholm 2008, Moloi & Chetty 2011). This informed the endeavour to ensure the workbooks compensated for the lack of PCK, and – as per the SACMEQ study – even basic content knowledge.

The second area was the emerging classroom-based research on the inadequate Opportunity to Learn (OTL) by Reeves and Muller (2005) and Reeves et al. (2008). Linked to this was the widespread acceptance that in at least 60% of primary schools in South Africa, there was a severe lack of basic text resources, especially textbooks (DBE 2011a; Howie et al. 2008; Moloi & Chetty 2011). The workbooks were thus conceptualised to alleviate these challenges with appropriate curriculum content and adequate curriculum pacing 'baked in' to the learner workbook, with the intention that the books would also be able to be taken home to extend learning time.

The third area focused on by the workbook team was the need for systematic phonics, which in 2010 when the DBE Workbooks were being conceptualised, was not the norm. Despite the growing influence of the UK's Rose Report (2006) which emphasised the need for systematic phonics instruction, there was only one study employing this approach at some level of scale in South Africa at the time, and that was Piper's (2009) Systematic Method for Reading Success (SMRS). It had been implemented in five provinces and the USAID evaluation showed promising results. The programme focused on systematically introducing letter sounds, blending sounds into words, recognising sight words and learning vocabulary while also supporting the development of comprehension skills. These findings guided the literacy approach underlying the development of the Grades 1-3 DBE Workbooks.

2.2.3 Review of international examples of workbooks

In addition to reviewing the South African evidence on learning outcomes and classroom practices, the workbook development team also explored international comparisons of similar interventions in other countries, where workbooks, resource-based and scaffolded teaching were used to improve learning outcomes at scale. Workbooks from the UK, specifically the SATS preparation workbooks and the Singaporean Mathematics and English worksheets, and scripted lessons from the United States were explored. Consideration was also given to workbooks used in France, Canada, Cuba, and New Zealand. These findings informed the instructional design and genre of the DBE materials that would act as a support tool to scaffold and guide teaching and learning – developed as a hybrid between a textbook and a workbook.

2.2.4 Assessing the classroom situation

In order to obtain first-hand experience of the resources being used in schools, classroom visits were conducted across a sample of 45 schools (across Eastern Cape, Gauteng, KwaZulu-Natal, Mpumalanga, and Limpopo) to obtain a baseline of Learning and Teaching Support Materials (LTSM) available and the methods used. Our findings confirmed the dire need for texts, posters, and charts, and for a more structured approach in the majority of classrooms (McKay 2013, 2018). Subsequently, in August 2010 the DBE convened a roundtable of experts, including teachers, to identify curriculum challenges that required attention. While the roundtable was valuable since it solicited the perspectives of teachers, it also initiated the participatory model used for the workbook development.

3 Development of the DBE Workbooks

Based on the research discussed above, the following decisions informed the development of the materials:

- To develop the materials in all 11 official languages
- To use a systematic phonics approach for initial literacy teaching and a language across the curriculum for the series
- To address the dearth of reading materials in classrooms
- To use the materials as a scaffold for teachers who required such support

Each of these decisions is discussed in the sections which follow.

3.1 Language of the workbooks

At the outset, the most important question facing the development team was what language(s) to use for the DBE Workbooks. Initially there was not a lot of support for the development of workbooks in all 11 official South African languages since this was still contested within the department in 2010, with some officials believing that a 'straight-for-English' approach would afford learners more opportunities to practise English, would be less cumbersome for development, printing and distribution, and less costly.

Notwithstanding this, the development of the materials proceeded in all 11 languages up to Grade 6, a hefty task involving considerable language and teaching innovation[5] and versioning across the 11 languages. The versioning process, as Sebate (2014, 3) explains, required the African language authors (1) to be familiar with the socio-cultural background of the learners, their ages, the appropriateness of the register as well as linguistic peculiarities; (2) to regard the source texts as a trigger for creativity and pedagogical artistry to incorporate the appropriate cultural and linguistic aspects, and (3) to version within the skopos or purpose of text and allowing for pedagogical adaptations. This meant that the materials were not translations from a single source

5 All the African language authors were PANSALB members who were able to engage their respective language boards for advice and to authenticate new or newly-coined words.

but that they were uniquely developed per language using the process developed for the workbooks.

Once the development had 'reached the point of no return', the decision for producing materials in 11 languages was accepted. This also contributed to reinforcing mother-tongue literacy beyond Grade 3. The prioritisation of African languages and 'mother-tongue-first' and/or bilingual/multilingual education had strong proponents both in South Africa (Alexander et al. 2007; Plüddemann 2010; MacDonald 1990, 1991, 2002) and internationally (UNESCO 2010a, 2010b; Alidou et al. 2006). UNESCO (2003) developed a normative framework, guidelines and principles for languages and education, making a strong case for mother-tongue instruction as a means of improving educational quality. Despite this minimum standard for quality education, Hanemann and McKay (2019, 353) point out that 40 per cent of the world's population do not access education in a language they speak or understand and that this was still contested terrain in South African policymaking circles in 2010. Since then the empirical evidence in support of the use of mother-tongue has grown considerably in South Africa (Taylor & Von Fintel 2016 and Hanemann & McKay 2019). Recent international reviews by RTI (2021) and the World Bank (2021) suggest there is universal consensus that children should learn to read and calculate in their home language (see also Erling et al. 2021).

3.2 Approach to Foundation Phase literacy

Having addressed the language question, it was necessary to decide on a research-based approach to guide the development of the literacy materials. At the time of development (2010/2011) South Africa was still steeped in a whole-language approach to reading from previous curricula (MacDonald 2002; Christie 2005; Hoadley 2018). With the introduction of the CAPS curriculum, the findings of the National Reading Panel (2000) in the US percolating through international research circles (Shanahan 2005; Piper 2009) and as with the UK's Rose Report, it was decided that the DBE Workbooks would follow a systematic phonics approach. This research further emphasised the importance of phonemic awareness, the alphabetic principle, and attention to reading fluency for improving reading comprehension (see Figure 2 below for an example worksheet from the Grade 1 Term 1 book teaching a letter sound). Over this period, during which these fundamental decisions were made, ongoing engagements were held with Helen Abadzi (World Bank), as well as with UNESCO colleagues such as Hasanna Alidou, Adama Ouane, and Ulrike Hanemann. There were also engagements with local literacy experts such as Vic Rodseth and Carol MacDonald.

In addition to using a systematic phonic approach, the texts (and related activities) were theme-based in an endeavour to teach language across the curriculum in view of enabling Cognitive Academic Language Proficiency (CALP) language and concepts needed for learning in other learning areas (Cummins 2008; McKay 2013).

Figure 2 Examples of phonics pages in the Grade 1 DBE Workbooks (isiZulu and English Term 1)

Source: isiZulu HL Grade 1 Term 1 and 2 Workbook, pages 40 and 38; English HL Grade 1 Term 1 and 2 Workbook, pages 30 and 32. From: https://www.education.gov.za/Curriculum/LearningandTeachingSupportMaterials(LTSM)/Workbooks/2021Workbook1.aspx

3.3 The dearth of reading materials in classrooms

The baseline research conducted by the workbook team in 2010 identified a dearth of reading materials as well as the the limitations teachers had in developing reading comprehension skills. It was decided that the new DBE Workbooks would include an extended reading text on every second worksheet, in line with the genre for reading required in the CAPS (such as folklore, instructional text, or a letter). In this way, it was possible to integrate 64 two-page reading texts for reading comprehension – providing immediate relief and compensation for the lack of reading materials. The aim was to increase the overall amount of reading in the classroom and help shift South African pedagogy away from only reading off the chalkboard, which was a common practice at the time. This finding was confirmed by Hoadley (2012) whose literature review of South African classroom-based research, found that "the most common form of reading consisted of the teacher writing three or four sentences on the board and students chorusing these after the teacher". Similarly, Taylor and Moyane (2004) found that in a study of 24 Limpopo primary schools, children interacted individually with books in only 3% of literacy classrooms.

3.4 Organising, sequencing, and pacing learning in the workbooks

Finally, the workbooks offered support for teachers who ordinarily would need to design and make copies of their own worksheets. They were conceptualised as a structured intervention for those teachers who experience challenges in providing activities; either because they lacked the resources or the ability to do so themselves. Importantly, the DBE Workbooks were designed to provide a minimum set of workbook activities for each child. They were not designed to be a complete programme of materials since that would require additional elements such as graded readers[6], big books, mathematics manipulatives, teachers' guides etc.

In order to ensure that the workbooks provide a strong scaffold, consideration was given to how they might organise learning. In response to this, the workbooks were sequenced and labelled and coloured by term and week in accordance with the CAPS, with each week comprising four worksheets. Apart from making the content, pacing and sequencing explicit to both teacher and learner, the workbooks also aimed to 'make learning visible', to facilitate curriculum coverage and provide support for teachers as they shifted from OBE to the CAPS.

Two additional practical considerations were taken into account: the first practical decision was to prepare lessons for only eight of the 10 weeks per term. It was recognised that schools rarely teach for a full 10 weeks due to a range of disruptions during the term, including teacher absenteeism, examinations, and other school activities. The 10 weeks of the CAPS was compressed into eight weeks of worksheets with four worksheets per week for literacy and four per week for mathematics. The second consideration was splitting each grade workbook into two volumes. In any one grade or subject[7] (literacy or numeracy) there were 128 worksheets (eight weeks x four terms x four worksheets/week = 128). Publishing all 128 worksheets in one book would result in a book that could not lie flat on a desk and which would be too cumbersome to carry. It was decided to split the worksheets into two volumes; Volume 1 for the first and second terms, and Volume 2 for the third and fourth terms (each with 64 worksheets). The table below provides an overview of the subjects per grade and year in which the workbooks were available[8].

6 It should be noted that in recent years (from 2014 to 2017) a series of DBE graded readers and DBE big books were developed for the Foundation Phase in all 11 languages and are available on the DBE website. They were distributed to lower-quintile schools and schools that were implementing the Incremental Introduction to African Languages (IIAL) policy.

7 The Life Skills books, first distributed in 2012, were also available in two volumes. However, since the CAPS for Life Skills was less extensive than for Home Language or Mathematics, there were only 32 worksheets per volume with two worksheets per week and eight weeks per term. The Grade R workbooks were published in four volumes, one per term. Each contained eight weeks' worth of worksheets that integrated literacy, numeracy and life skills.

8 Our focus in this chapter is the literacy and numeracy workbooks. As the table shows, in 2012 the brief expanded significantly to include Life Skills for Grades 1-3 in all 11 official languages and that English as a First Additional Language (EFAL) be introduced from Grade 1 in 2012, the conceptual focus of which was additive bilingualism. Similarly, in 2016 the brief was further extended to develop workbooks to support the Incremental Introduction of African Languages (IIAL) to promote previously marginalised languages and promote social cohesion. These are listed as 'SAL', Second Additional Language, in the table.

Table 1 Overview of grades, subjects and languages by year (number in cell is the number of languages this book is available in)

Grade	Subject	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
Grade R	HL+Num+LS		11	11	11	11	11	11	11	11	11	11
Grade 1-3	Home Lang	11	11	11	11	11	11	11	11	11	11	11
	Numeracy	11	11	11	11	11	11	11	11	11	11	11
	Life Skills		11	11	11	11	11	11	11	11	11	11
	English First Additional Language		Eng	Eng	Eng	Eng	Eng	Eng	Eng	Eng	Eng	Eng
	Second Additional Language*						9	9	9	9	9	9
Grade 4-6	Home Lang		11	11	11	11	11	11	11	11	11	11
	Numeracy		Eng/Afr	Eng/Afr	Eng/Afr	Eng/Afr	Eng/Afr	Eng/Afr	Eng/Afr	Eng/Afr	Eng/Afr	Eng/Afr
	EFAL		Eng	Eng	Eng	Eng	Eng	Eng	Eng	Eng	Eng	Eng
Grade 7-9	Home Lang		11	11	11	11	11	11	11	11	11	11
	Numeracy		Eng/Afr	Eng/Afr	Eng/Afr	Eng/Afr	Eng/Afr	Eng/Afr	Eng/Afr	Eng/Afr	Eng/Afr	Eng/Afr

* *Second Additional Language excludes English and Afrikaans. Although the PDFs are available on the DBE website, they have not been printed and distributed up to 2021*

3.5 The team of experts involved in developing the workbooks

Given the vast range of materials needed for the DBE Workbook project, the development team included language practitioners, curriculum experts, instructional designers, layout designers, photographic designers, children's artists, technical artists, editors, and translators. The artists, layout designers, and editors worked with the authors in conceptualising the pedagogical design. A detailed description of the development team, the criteria for their selection, and a list of their skills is available on request from the authors.

The original team comprised the same team of language, literacy, and mathematics experts who had been engaged in the Kha Ri Gude materials development. The team of African language literacy and language experts who had been engaged in the development of the KRG learner workbooks were drawn from the universities. Additional language and mathematics experts were approached from NGOs. They were contracted on an hourly basis against staff vacancies within the DBE. This meant that their appointments were constantly being rotated as the vacancies were filled. In 2010 the full literacy and numeracy team consisted of approximately 48 people (full-time equivalent) at peak times. Apart from Veronica McKay, the entire workbook development team was external to the DBE. Additional existing KRG staff who were officials in the DBE[9] took on additional functions to support the process. The development team was contracted for approximately one month per grade per book (120-160 hours). The group of artists worked according to artwork briefs and typesetters and instructional designers were allocated a number of books needed to be typeset per month.

Since a sequential development process was used, the team was recycled with books being versioned, piloted and proofread, and corrected simultaneously with the development of new books.

In brief, materials developers were required to follow the CAPS including the assessment guidelines, to design lessons in accordance with the templates, thematic approaches and activities, as well as to be engaged in the review, revision, and piloting the books.

3.6 Developing familiarity with assessment formats

An additional concern of the materials development team was exposing children to different assessment styles, especially multiple choice questions (MCQ), which were not widely used, most notably in no-fee schools, at the time. This is understandable given that MCQs are typically an individualised assessment and cannot really be done on a blackboard. To that end, MCQs were incorporated into the workbooks. Use was also made of the released passages and assessment types from international assessments such as PIRLS, pre-PIRLS, and the British SATs that were available at that stage. These were adapted and versioned to give learners exposure to these formats. See for example

9 Prof. Veronica McKay was responsible for the project, including the development of the concept and managing the various teams and she wrote a number of workbooks. Dr Obert Maguvhe provided the direction for LSEN and the Braille texts, Marietta du Toit took on the role of Finance Officer, and Elijah Sekgobela managed the teams of fieldworkers involved in the pilots.

the non-fiction text and assessment "A school trip for insect lovers" in the Grade 4 literacy workbook (DBE 2021, 24).

4 Existing literature on the workbooks

A survey of the literature on South African schooling over the last three decades reveals that there have been surprisingly few policy interventions that have actually changed the grammar of schooling in the national schooling system. While there have been some bold provincial and regional approaches to changing teaching and learning in primary schools (for example GPLMS, PILO, EGRS, LitNum, and NECT[10]) none of these programmes have been implemented country-wide, and even many of these well thought-out interventions lack evidence to suggest that they have changed the ways that teaching and learning happens in South African classrooms. By contrast, some scholars have argued that overall improvements in educational outcomes over the 2010-2019 period are largely attributable to national policy changes rather than interventions, emphasising a new curriculum, new assessments and new materials (Gustafsson, 2019; 2020).

Of these three country-wide reforms – CAPS, ANAs, and the workbooks – the first and second have received much more scholarly attention than the last. The South African curriculum shift from the 'outcomes-based' Curriculum 2005 to the 'knowledge-centric' Curriculum and Assessment Policy Statement (CAPS) is well-documented (Jansen & Christie, 1999; Pinar, 2010; Hoadley 2018; Hoadley & Muller 2019). To a lesser extent, the implementation of the nation-wide language and mathematics assessments in the form of the Annual National Assessments (ANAs) from 2011 to 2014 has also been well-documented in reports by the Department of Basic Education itself (DBE, 2014) as well as in the scholarly literature (Graven & Venkat 2014; Van der Berg 2015; Thulare 2018; Gustafsson 2019; Unisa 2014; SAIDE 2012).

By contrast, the scholarly attention paid to the DBE Workbooks is minimal. Apart from a DBE-sponsored review by the Australian Council of Education Research (ACER 2013), subsequently published as Outhred et al. (2014), there are only ten peer-reviewed articles focusing on the DBE Workbooks in the last decade, and in most instances on hyper-specialised areas of interest. In chronological order the ten articles are as follows:

- An early study conducted by SAIDE/Zenex on the utility of the workbooks in classrooms (SAIDE 2012)
- A case study of how six Grade 3 mathematics teachers use the DBE Workbooks in their teaching (Matthews et al. 2014)
- An analysis of the Grade 3 language and mathematics DBE Workbooks as curriculum tools (Hoadley & Galant 2016)
- A sociological analysis of the pictures used in the mathematics DBE Workbooks (Pausigere 2017)

10 These are the Gauteng Primary Language and Mathematics Strategy (GPLMS) in Gauteng (Fleisch & Schšer 2014), the Programme to Improve Learning Outcomes (PILO) in KwaZulu-Natal (Metcalfe & Witten, 2019), the Early Grade Reading Study (EGRS) (Cilliers et al. 2019), the Western Cape's LitNum Project (JET 2013), and the National Education and Collaboration Trust (NECT).

- An analysis of the social and environmental values embedded in the DBE Workbooks (McKay 2018)
- A review of additive relations in the Grade 1-3 DBE Workbooks (Mostert 2019)
- A review of pedagogical and content support provided by the DBE Workbooks, drawing on different reviews of the workbooks (McKay 2019)
- An investigation into the alignment between content standards and workbook activities on numeric and geometric patterns in Senior Phase mathematics (Qhibi et al. 2020)
- An investigation into the alignment between senior phase mathematics content standards and numeric and geometric patterns' workbook activities (Dulu (2019).
- An overview of how the DBE Workbooks teach sustainability through 'multispecies relationality' and Ubuntu in the Life Skills workbooks (Mabunda & McKay 2021)

In addition to the above there was one commissioned study on the mechanics of the workbooks requested by AusAid (McKay 2013) and a few master's theses based on document-analyses that look at, for example, the values embedded in the workbooks (Rawhani, 2013). Given the ubiquity of the DBE Workbooks, it is unsurprising that they have been included in studies whose primary interest is not the workbooks per se, but rather teaching and learning in rural schools (NEEDU 2014) or as part of evaluations of reading and mathematics interventions (Ardington & Henry 2021).

The most cited of the peer-reviewed studies (Hoadley & Galant 2016) concludes that there is "moderate-to-strong" alignment between the workbooks and the curriculum and "in line with the DBE's intended purpose for the workbooks, [they] best represent a practice curriculum tool" rather than as the sole teaching or assessment tool. This is something that teachers themselves seem to agree with, as discussed in the next section.

5 Data

Given the ubiquity of the DBE Workbooks, it has become commonplace to include questions on the workbooks in large-scale government-led surveys. Previous research has validated that the DBE Workbooks are in fact available in virtually every school. For example, in the 2018 General Household Survey Statistics South Africa surveyed 21,908 households and asked respondents whether Grade 1-9 children in the household had access to the DBE Language and Mathematics Workbooks. The survey showed that 95-98% of respondents in every province reported that their child did have access to both a language and mathematics workbook (DBE 2018b, 22).

In the next section, we present the findings from a large-scale nationally-representative survey of South African schools conducted in 2017; the School Monitoring Survey (SMS) which included approximately 1,000 primary schools and 1,000 high schools. The 2017 SMS survey included a module specifically on the DBE Workbooks, making it especially helpful for the present analysis. We focus here on the primary school sample and specifically on the responses from the 1,543 Grade 3 teachers surveyed. We selected Grade 3 for two reasons: firstly, the focus of the DBE Workbooks was initially literacy and numeracy, which is the focus of the Foundation Phase (Grades

R-3), and secondly since the DBE Workbooks cover all of the four subject areas of home language, life skills, mathematics, and English first additional language, and at the Grade-3 level these are all taught by the same teacher.

6 Analysis and discussion

6.1 School Monitoring Survey 2017

Table 2 below reports the findings from the nine main questions that make up the Workbooks Module. All responses were mutually exclusive multiple choice responses. The responses of the Grade 3 teachers accord with the general perception in South African research circles that the workbooks are ubiquitous and frequently used. Nearly two-thirds of teachers reported that they use the DBE Workbooks “always” (63%) or “often” (28%). Put differently, more than 90% of teachers use it always or often, with minimal variation between responses for language workbooks and mathematics workbooks.

When asked to indicate the percentage of class time where the teacher used the DBE Workbooks, a third (32%) reported “100%” with a further third (36%) reporting either 66% or 75%, i.e. at least two-thirds of teachers (68%) are reporting that they use the DBE Workbooks for at least two-thirds (66%+) of their class time. Four in five teachers (81%) reported that learners use the workbooks both in class and at home and when asked specifically whether the teacher used the DBE Workbook for homework, 60% reported that they did and 77% reported that they use it for classwork.

Teachers were also asked a series of questions about how satisfied they were with the DBE Workbooks and to what extent the books were helpful in their teaching. Most teachers (79%) were either “satisfied” or “very satisfied” with the quality of the workbooks and about the same percentage (84%) reported that they found the DBE Workbooks “helpful” (42%) or “very helpful” (42%).

The final set of questions in Table 2 reports on whether teachers perceive the DBE Workbooks to be “sufficient on their own, one does not need a textbook”, with a similar question asking whether a textbook is sufficient on its own and does not need a workbook. Responses show that about 70% of teachers disagree that a workbook is sufficient on its own (i.e. without a textbook) and about 85% of teachers disagree that a textbook is sufficient on its own (i.e. without a workbook). Again there is little variation across language and mathematics. These responses coincide with Hoadley and Galant’s (2016, 11) analysis of what is in the Grade 3 workbooks, and who argue that by themselves the DBE Workbooks are not enough to cover teaching, assessment, and practice. They explain that

> *In order to use the workbooks to support quality teaching they would need further development to make the evaluative criteria and conceptual signalling in the text more explicit. This could be accomplished by strengthening the teacher notes, providing more comprehensive in-text notes, producing an aligned teacher guide, or using the workbook alongside a good textbook. With the recent proposal to produce a single textbook per subject per grade, this textbook could usefully be aligned with*

> *the workbooks. The textbook could then function as a primary transmission text, with clear conceptual signalling as well as relevant tasks, and the workbook could function as a practice tool, either for use in class or as a homework resource*
>
> Hoadley & Galant (2016, 11).

As the School Monitoring Survey (DBE 2017) shows, 69% of the teachers indicate that they use textbooks, which they (or their HODs) select from the school catalogue. Given this high number, we agree with Hoadley and Galant who state that the workbooks and textbooks should be aligned so as to offer further support for teachers with poor content knowledge.

Table 2 Reported use of DBE Workbooks and textbooks by Grade 3 teachers in South Africa (School Monitoring Survey 2017, *n*=1,543 teachers)

Question...	%	SE	%	SE	%	SE	%	SE
How often do you use the DBE workbook in your class?	Never		Occasionally		Often		Always	
Maths	**0%**	0,2%	**8%**	0,8%	**28%**	1,4%	**63%**	1,5%
Language	**0%**	0,1%	**8%**	0,9%	**30%**	1,4%	**62%**	1,5%
Please indicate the % of class time you have used the DBE workbook?	0%, 25% or 33%		50%		66% or 75%		100%	
Maths	**10%**	0,9%	**22%**	1,3%	**36%**	1,5%	**32%**	1,4%
Language	**10%**	0,9%	**20%**	1,3%	**37%**	1,5%	**33%**	1,4%
Indicate the main purposes for which you use the DBE workbooks in your class?	For classwork		For homework		Selected examples for demonstration during lessons		Other	
Maths	**77%**	1,3%	**60%**	1,6%	**35%**	1,5%	**9%**	0,9%
Language	**82%**	1,2%	**60%**	1,6%	**33%**	1,5%	**8%**	0,9%
Do learners use the DBE workbooks in class only, at home only, or both?	In class only		At home only		In class and at home			
Maths	**17%**	1,2%	**2%**	0,5%	**81%**	1,2%		
Language	**19%**	1,3%	**1%**	0,4%	**79%**	1,3%		
How satisfied are you with the quality of the DBE workbook?	Not satisfied		Somewhat satisifed		Satisfied		Very satisfied	
Maths	**6%**	0,7%	**16%**	1,1%	**47%**	1,5%	**32%**	1,4%
Language	**5%**	0,6%	**12%**	0,9%	**51%**	1,5%	**33%**	1,4%
How helpful have you found the DBE workbooks in your work?	Not helpful		Somewhat helpful		Helpful		Very helpful	
Maths	**3%**	0,5%	**14%**	1,0%	**42%**	1,5%	**42%**	1,5%
Language	**2%**	0,4%	**12%**	1,0%	**45%**	1,5%	**40%**	1,5%
The DBE workbooks are sufficient on their own, one does not need a textbook	Strongly Agree		Agree		Disagree		Strongly Disagree	
Maths	**9%**	0,8%	**26%**	1,3%	**51%**	1,5%	**15%**	1,1%
Language	**7%**	0,7%	**20%**	1,2%	**59%**	1,5%	**14%**	1,1%
Is a textbook being used to teach _______ in this class?	Yes		No					
Mathematics	**68%**	1,4%	**32%**	1,4%				
Language	**69%**	1,4%	**31%**	1,4%				

Question...	%	SE	%	SE	%	SE	%	SE
Textbooks are sufficient on their own, one does not need to use the DBE workbooks	Strongly Agree		Agree		Disagree		Strongly Disagree	
Maths	**3%**	0,5%	**13%**	1,0%	**65%**	1,4%	**19%**	1,2%
Language	**3%**	0,5%	**13%**	1,0%	**67%**	1,4%	**18%**	1,2%

6.2 Factors that led to the successful development, roll-out, and use of the DBE Workbooks

The DBE Workbooks are the only universally available resource in every school. The 2017 School Monitoring Survey classroom audits conducted in a sample of approximately 2,000 schools revealed that more than 95% of surveyed learners had both their literacy and numeracy workbooks (DBE, 2018a: 55). Teachers are clearly satisfied with the quality of the workbooks and use them extensively. Teachers state that they are well-aligned to the curriculum and offer pacing and sequencing guidance. Given all of this, it is worthwhile asking: what has made the workbooks so enduring? We propose the following:

1. **Political continuity:** The DBE Workbooks were one of the first major initiatives overseen by minister of Basic Education, Ms Angie Motshekga in 2010. Her three terms as minister have provided political stability for the project.
2. **A 'universal national approach' rather than an 'intervention approach':** Much of South Africa's education reform landscape in 2022 is dominated by small- to medium-scale local interventions. By contrast, the workbooks were rolled out to all 24,000 schools creating a new 'minimum floor' of text resources in every school. Because they are universally available and nationally endorsed, teachers use them, subject advisors theoretically monitor them, and parents have some window into their children's world of school. Before the DBE Workbooks many schools in poor and rural communities had no books and mainly worked off the blackboard. Providing six to eight full-colour workbooks to every child each semester was a big step forward for text provision in the country.
3. **Alignment to simultaneous national curriculum and the changes in assessments:** The fact that the policies for the CAPS and the ANAs coalesced with the introduction of the workbooks, all being implemented in the same year (2011) made the success of all three interventions much more likely. The sense that there was a real national education reform was widespread at that time.
4. **Realistic content level:** The workbook development team was realistic about current levels of teacher content-knowledge, prevalent class sizes, other materials that were present, and the lack of preparation and training of most teachers. This informed the simple one-book-per-subject-per-child-per-year approach, with logical sequencing and activities throughout the series.
5. **Cost-effective development, printing, and delivery:** The fact that the books are royalty-free and incur only printing and delivery costs radically reduces the cost of providing the books to schools. Economies of scale led to a small number of large print runs and a small number of large distribution contracts.

6.3 Plausible contribution to the national improvement in learning outcomes

In this chapter we have not presented evidence that the DBE Workbooks have led to an increase in learning outcomes among children as there have not been any quantitative studies that measure the causal impact of the DBE Workbooks on children's learning outcomes. We recognise that no single intervention can change the over-arching forms of pedagogy and assessment in the country.

Notwithstanding the above, we would agree with other researchers (Gustafsson 2019) that the universal distribution and use of the DBE Workbooks is one of the main contributing factors to improvements in overall learning outcomes seen across a number of independent studies such as PIRLS, TIMSS, and SACMEQ (Van der Berg & Gustafsson 2019). Clearly, there have been changes in what children learn and are able to do, and it is reasonable to expect that one of the largest three reforms in the last decade should take some credit for those changes. One could even argue that much of the new curriculum (CAPS) was made manifest to teachers largely through the workbooks, which are the most practical instantiation of the curriculum teachers are likely to encounter. Now that the ANAs no longer exist (and haven't since 2014), one of the only examples of proper grade-level assessments at the primary school level are the DBE Workbooks, together with provincial common exams (for example, as used in the Western Cape).

7 Conclusion

Changing what teachers do in the classroom is notoriously difficult. That includes what they teach, how they teach, how much they teach, and what they use to teach. While the DBE Workbooks have not changed all of these (and may have changed some more than others) it seems undeniable that the DBE Workbooks have percolated through South African schools to such an extent that they are now the only teaching tool that an observer is guaranteed to see in use in South African classrooms. Whether or not they were designed to be the core teaching tool of the curriculum, this is how teachers are now using them, in particularly in most no-fee schools. They are provided every year without fail and offer an element of structure and sequencing that was previously lacking.

Like all resources there are elements of the DBE Workbooks that can and must be improved over time. Yet it is clear to us that building upon existing successes is much more likely to yield positive results than taking a new approach.

Suggestions for a way forward include, for example, creating subject-and-grade-specific teacher guides[11] and/or a textbook that is explicitly aligned to the DBE Workbooks. The introduction of an additional minimum set of literacy and numeracy resources could be explicitly aligned to (and integrated into) the language and mathematics workbooks and accompanying guides/textbooks. These might include an anthology of graded readers, classroom libraries, mathematics manipulatives etc. This

11 Many teacher guides were developed as part of the project. For cost reasons, they were not published.

could mark a second era in the provision of universal resources for the improvement of literacy and numeracy in South Africa.

In this chapter we presented an overview of both the genesis and development of the DBE Workbooks. We showed how political, financial, and administrative factors contributed to the success of the intervention. By taking a universal approach the DBE Workbooks have essentially created a new 'minimum floor' of text resources in all South African classrooms, and one that has arguably changed what teachers use in their teaching and, possibly, also what and how they teach. Given the high use and teacher satisfaction shown by the nationally representative data, it is not an overstatement to say that the DBE Workbooks have become the de facto curriculum tool in South Africa, especially in no-fee schools. It is for this reason that we think there is a strong argument to be made that the roll-out of the DBE Workbooks to all schools since 2011 has played an important role in contributing to the improvement in learning outcomes over the last decade. In line with Hoadley's (2018, 72) argument, a "pedagogic device is not a thing, but rather, like grammar, it is a set of rules that operates in the background, generative of particular processes and forms, in this case the process of knowledge becoming pedagogic communication". We would argue that given their ubiquitous presence and universal use, the DBE Workbooks are a rare example of resource-led changes to how knowledge becomes pedagogic communication, and possibly to an enduring change to the grammar of South African schooling.

References

Alexander, N., Bamgbose, A., Bloch, C., Busch, B., Coste, D., Edwards, V., Fal, A., Kum'a Ndumbe III, P., Samessékou, A. (2007). Literacy and linguistic diversity in a global perspective. European Centre for Modern Languages. Available at: http://www.praesa.org.za/wp-content/uploads/2018/07/COE_-Linguistic-diversity-busch07_LDL_E.pdf (Accessed 2 November 2021).

Alidou, H., A. Boly, B. Brock-Utne, Y. Satina Diallo, K. Heugh, and H. Ekkehard Wolff. (2006). Optimizing Learning and Education in Africa – The Language Factor: A Stock-Taking Research on Mother Tongue and Bilingual Education in Sub-Saharan Africa. Paper presented for discussion at the ADEA Biennial Meeting, Libreville, Gabon, March 27–31. Available at: http://www.adeanet.org/adea/downloadcenter/Ouga/B3_1_MTBLE_en.pdf (Accessed 9 November 2018).

Ardington, C. and Henry, J. (2021). Funda Wande Limpopo Evaluation: Midline Report. Available at: https://fundawande.org/img/cms/news/Limpopo%20TA%20Evaluation%20 2021.pdf (Accessed 21 February 2022).

Australian Council for Educational Research (ACER). (2013). Formative evaluation of textbooks and workbooks in South Africa. DBE, Pretoria.

Carnoy, M., & Chisholm, L. (2008). Towards understanding student academic performance in South Africa: A pilot study of Grade 6 mathematics lessons in Gauteng province. Pretoria: HSRC

Christie, F. (2005). *Language Education in the Primary Years*. Sydney: NSW Press.

Cilliers, J., Fleisch, B., Prinsloo, C. & Taylor, S, (2019). How to improve teaching practice? An experimental comparison of centralized training and in-classroom coaching. *Journal of Human Resources*. 0618–9538R1.

Cummins J. (2008). BICS and CALP: Empirical and theoretical status of the distinction. In: Hornberger, N.H. (eds) *Encyclopedia of Language and Education*. Boston, MA: Springer. https://doi.org/10.1007/978-0-387-30424-3_36

DBE (Department of Basic Education). (2010). Proposal for the Development of the Workbooks. Revised proposal. Pretoria: DBE.

DBE (2011). Annual National Assessments 2011: A Guideline for the Interpretation and Use of ANA Results. Pretoria: DBE. Available at: https://edulibpretoria.files. wordpress.com/2008/01/ana-2011-guidelines-for-interpretation.pdf. (Accessed 9 November 2018).

DBE (2014). Report on the Annual National Assessments of 2014. Department of Basic Education. (Online). Available at: https://www.education.gov.za/Portals/0/Documents/Reports/REPORT%20ON%20THE%20ANA%20OF%202014.pdf?ver=2014-12-04-104938-000 (Accessed 19 April 2021).

DBE (2018a). The 2017 School Monitoring Survey: Quantitative Survey Main Report. Department of Basic Education. Online. Available at: https://www.education.gov.za/Research,Monitoring EvaluationReports.aspx (Accessed 29 April 2021).

DBE (2018b). General Household Survey: Focus on Schooling 2018. Department of Basic Education. (Online). Available at: https://www.education.gov.za/Portals/0/Documents/Reports/20191104_GHS_final.pdf?ver=2019-11-04-111833-313 (Accessed 29 March 2022).

DBE. (2021). Grade 4 English Home Language Workbook 1 (Terms 1 & 2). Available at: https://www.education.gov.za/Portals/0/Documents/Manuals/2021%20Workbooks/WB1/HL/ENG/HL_ENG_Gr4_B1.pdf?ver=2021-02-01-102321-910 (Accessed 28 March 2022).

DoE. (Department of Education). (2008). Grade 3 Systemic Evaluation 2007 Leaflet. Pretoria: DoE.

DoE (2010). Intensifying Literacy and Numeracy Support to Schools. Workbook Proposal, 21 January, 2010.

Dulu, Q. (2019). Alignment between senior phase mathematics content standards and numeric and geometric patterns' workbook activities. Masters Thesis. University of Limpopo. Available at: http://ulspace.ul.ac.za/bitstream/handle/10386/3191/qhibi_ad_2019.pdf?sequence=1&isAllowed=y (Accessed 19 April 2021).

Erling, E., Clegg, J., Rubagumya, C., & Reily, C. (2021). *Multilingual Learning and Language Supportive Pedagogies in Sub-Saharan Africa*. Routledge: Oxfordshire.

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

Fleisch, B. & Schöer, V. (2014). Large-scale instructional reform in the Global South. *South African Journal of Education*. 23 (4).

Graven, M., & Venkat, H. (2014). Primary Teachers' Experiences Relating to the Administration Processes of High-stakes Testing: The Case of Mathematics Annual National Assessments. *African Journal of Research in Mathematics, Science and Technology Education*. 18.

Gustafsson, M. (2019). Pursuing equity through policy in the schooling sector 2007-2017. In Spaull, N. & Jansen, J. (Eds) *South African Schooling: The Enigma of inequality*. Cham: Springer.

Gustafsson, M. (2020). A revised PIRLS 2011 to 2016 trend for South Africa and the importance of analysing the underlying microdata. Stellenbosch Economic Working Papers WP02/2020. Stellenbosch.

Hanemann, U. & McKay, V. (2019). Learning in the mother tongue: Examining the learning outcomes of the South African Kha Ri Gude literacy campaign. *International Review of Education*. 65, 351–387.

Hoadley, U. (2012). What do we know about teaching and learning in primary schools in South Africa? A review of the classroom-based research literature. *Education as change*, 16 (2) 187-202.

Hoadley, U. (2018). *Pedagogy in Poverty*. Abingdon: Routledge.

Hoadley, U., & Galant, J. (2016). An analysis of the Grade 3 Department of Basic Education workbooks as curriculum tools. *South African Journal of Childhood Education*, 6(1), 1–12.

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. University of Pretoria: Centre for Evaluation and Assessment.

Jansen, J. and Christie, P. (1999). *Changing Curriculum: Studies on Outcomes-based Education in South Africa*. Cape Town: Juta.

JET Education Services. (2013). Consolidated evaluation report of the WCED literacy and numeracy intervention project: Final report. Unpublished.

Matthews, C., Mdluli, M., & Ramsingh, V. (2014). The use of workbooks in South African Grade 3 mathematics classrooms. *South African Journal of Childhood Education*. 4(1).

Mabunda, P. and McKay, V. (2021). Education curriculum and multispecies relations. In: McIntyre-Mills, J. & Corcoran-Nantes, C. (Eds). *From Polarisation to Multispecies Relationships*. Cham: Springer.

MacDonald, C. (1990). *Crossing the Threshold into Standard Three in Black Education: The Consolidated Main Report of the Threshold Project*. Pretoria: HSRC Press.

MacDonald, C. (1991). *Eager to Talk and Learn and Think*. Cape Town: Maskew Miller Longman.

MacDonald, C. (2002). Are the children still swimming up the waterfall? A look at literacy development in the new curriculum. *Language Matters*, 33(1), 111–141.

McKay, V. I. (2013). *The Development of the South African School Workbooks*. Canberra: AusAID Education Resource Facility.

McKay V. (2018). Introducing a parallel curriculum to enhance social and environmental awareness in South African school workbooks. In: McIntyre-Mills J., Romm N., Corcoran-Nantes, Y. (eds) *Balancing Individualism and Collectivism. Contemporary Systems Thinking.* Springer, Cham.

McKay, V. (2019). Fit for purpose: Using a distance education approach to support underperforming schools in South Africa. *Progressio*. 41 (1).

McKay, V. (2020). Learning for development: Learners' perceptions of the impact of the Kha Ri Gude Literacy Campaign. *World Development*, 25, 104684.

Metcalfe, M. and Witten, A. (2019). Taking change to scale: Lessons from the Jika iMfundo campaign for overcoming inequalities in schools. In Spaull, N. & Jansen, J. (Eds): *South African Schooling: The Enigma of Inequality*. Cham: Springer

Moloi, M. & Chetty, M. (2011). The SACMEQ III Project in South Africa: A Study of the Conditions of Schooling and the Quality of Education – South Africa. SACMEQ. Available at: https://www.gov.za/documents/sacmeq-iii-project-south-africa-study-conditions-schooling-and-quality-education (Accessed 15 November 2022).

Muller, J., & Hoadley, U. (2019). Curriculum reform and learner performance: An obstinate paradox in the quest for equality. In: Spaull, N. & Jansen, J. (Eds) *South African Schooling: The Enigma of Inequality*. Cham: Springer. 109-125.

NEEDU. (2014). NEEDU National Report 2013: Teaching and Learning in Rural Primary Schools. National Education Evaluation Development Unit. Department of Basic Education. Pretoria.

Outhred, R., Deliwe, C.N., Stubberfield, C., Beavis, A., Wilkinson, J., Murphy, M. (2014). Towards quality as an equity imperative. In: Zhang, H., Chan, P.W.K., Boyle, C. (Eds) *Equality in Education*. Rotterdam: SensePublishers.

Pausigere, P. (2017). A sociological analysis of the pictures used in the South African foundation phase numeracy workbooks and their mathematical implications. *Journal of Education*. 70.

Pinar, W. F. (2010). Curriculum studies in South Africa: Intellectual histories & present circumstances. London: Palgrave Macmillan.

Piper, B. (2009). Integrated Education Program Impact Study of the SMRS. USAID. (Online). Available at: https://nicspaull.files.wordpress.com/2015/01/piper-2009-egra-south-africa-smrs.pdf (Accessed 25 October 2021).

Plüddemann, P. (2010). Home-language based bilingual education: Towards a learner-centred language typology of primary schools in South Africa. PRAESA Occasional Papers No. 32. (Online). Available at: http://www.praesa.org.za/wp-content/uploads/2016/09/Paper32-1.pdf (Accessed 2 November 2021).

Qhibi, A. D, Dhlamini, Z. B. and Chuene, K. (2020). Investigating the strength of alignment between Senior Phase mathematics content standards and workbook activities on number patterns. *Pythagoras*, 41(1).

Rawhani, V. (2014). An analysis of the values in the Foundation Phase of the South African Curriculum and their application in the work books. Masters thesis. University of KwaZulu-Natal. (Online). Available at: https://ukzn-dspace.ukzn.ac.za/bitstream/handle/10413/10784/Rawhani_Veda_Chowghi-Abd-el_Fattah_2013.pdf?sequence=1&isAllowed=y (Accessed: 19 April 2021).

Reeves, C. and Muller, J. (2005). Picking up the Pace: Variation in the Structure and Organization of Learning School Mathematics. *Journal of Education*, 37(1), 103–130.

Reeves, C., Heugh, K., Prinsloo, C.H., Macdonald, C., Netshitangani, T., Alidou, H., Diedericks, G. & Herbst, D. (2008). Evaluation of Literacy Teaching in Primary Schools of Limpopo Province. Commissioned by the Limpopo Department of Education (LDoE). HSRC in Association with the Department of Language Education at the University of Limpopo.

RTI. (2021). Language of Instruction in Foundational Literacy & Numeracy Programs in sub-Saharan Africa: The Basics. Science of Teaching. (Online). Available at: https://scienceofteaching.s3.eu-west-3.amazonaws.com/assets/7JE4Li4fWGMnEVNg_LM-jl18igTgB1rJq-Language%20of%20Instruction%20Basics.pdf (Accessed 2 November 2021).

Saide (South African Institute for Distance Education). (2012). Rapid Research Assessment of the Use of Department of Basic Education Workbooks. http://www.saide.org.za/resources/newsletters/Vol_18_no.2_2012/Content/Zenex%20Rapid%20Research.htm (Accessed 9 November 2018).

Sebate, P. M. (2014). Considering the socio-cultural background in the translation of Grade 1–6 literacy and numeracy workbooks in South Africa. Paper presented at the 34th International Society for Teacher Education (ISfTE) Conference, 22–25 April in Antalya, Turkey. Unpublished.

Shanahan, T. (2005). The National Reading Panel Report: Practical Advice for Teachers. (Online). Available at: https://files.eric.ed.gov/fulltext/ED489535.pdf (Accessed 2 November 2021).

South African Parliament. (2008). Government Notice: Department of Education Foundations for Learning Campaign. Government Gazette No. 30880 (306). (Online). Available at: https://www.eccurriculum.co.za/FoundationPhase/FFL-Gazette30880-2008.pdf (Accessed 4 October 2021).

Taylor, N. (2011). National School Effectiveness Study - Synthesis Report. Johannesburg: JET Education Services.

Taylor, N. & Moyane, J. (2004). Khanyisa Education Support Programme: Baseline study part 1: communities, schools and classrooms. Johannesburg: JET Education Services.

Taylor, S. & Von Fintel, M. (2016). Estimating the impact of language of instruction in South African primary schools: a Fixed effects approach. *Economics of Education Review* 50(75–89).

Thulare, T.D. (2018). A Policy Analysis of the Annual National Assessments in South Africa, in Wiseman, A.W. and Davidson, P.M. (Eds) *Cross-nationally Comparative, Evidence-based*

Educational Policymaking and Reform: International Perspectives on Education and Society, Vol. 35, Bingley: Emerald Publishing Limited, 71–100.

Tyack, D. and Tobin, W. (1994). The "Grammar" of Schooling: Why Has it Been so Hard to Change? in *American Educational Research Journal*. 31 (3), 453–479.

Van der Berg, S. (2015). What the Annual National Assessments can tell us about learning deficits over the education system and the school career. *South African Journal of Childhood Education*. 5(2).

Van der Berg, S., & Gustafsson, M. (2019). Educational outcomes in post-apartheid South Africa: Signs of progress despite great inequality. In N. Spaull & J. Jansen (Eds), *South African schooling: The enigma of inequality*. New York: Springer Nature.

UNESCO. (2010a). Policy Guide on the Integration of African Languages and Cultures into Education Systems. UNESCO. (Online). Available at: https://uil.unesco.org/fileadmin/keydocuments/Africa/en/policy_guide_sep_web_en.pdf (Accessed 2 November 2021).

UNESCO. (2010b). Why and how Africa should invest in African languages and multilingual education. (Online). Available at: https://uil.unesco.org/literacy/multilingual-research/why-and-how-africa-should-invest-african-languages-and-multilingual (Accessed 2 November 2021).

University of South Africa (Unisa) (2014). *The 500 schools project: Making schools better*. Pretoria: Unisa.

World Bank. (2021). Loud & Clear: Effective Language of Instruction Policies for Learning. World Bank. (Online). Available at: https://documents1.worldbank.org/curated/en/517851626203470278/pdf/Loud-and-Clear-Effective-Language-of-Instruction-Policies-For-Learning.pdf (Accessed 2 November 2021).

Zuma, J. (2010). State of the Nation Address (11 February 2010). Parliament. (Online). Available at: https://www.sahistory.org.za/archive/2010-president-zuma-state-nation-address-11-february-2010-0 (Accessed 8 October 2021).

03

Coaching research in the Early Grade Reading Studies in South Africa

BRAHM FLEISCH & KAITLIN ALSOFROM

Abstract

Over the past two decades, instructional coaching has emerged as a key feature of effective interventions to improve early grade learning. This chapter traces South African research undertaken over the past twelve years on the impact of instructional coaching in large-scale structured pedagogic programmes. Combining evidence from large-scale counterfactual comparable and replication studies and in-depth qualitative case studies that aim to uncover mechanisms, this chapter illustrates the contribution of this research programme to advancing knowledge of how to improve early grade learning system-wide in the Global South. Specifically, the research highlights the feasibility, cost-effectiveness and mechanisms of on-site instructional coaching as part of structured pedagogic programmes.

KEYWORDS
EGRS, instructional coaching, structured pedagogy, randomised control trials

1 Introduction

This chapter provides a narrative description and analysis of the last two decades of South African research on instructional coaching as part of structured pedagogic programmes. The narrative approach of this article details, chronologically, the impact of coaching on early grade reading using various counterfactual research designs and methodologies that complement each other – bringing into focus the importance of coaching for change in the South African context. The aim of this article is to highlight important findings and signal new and important insights gained in South Africa through tracking a range of studies which have included coaching as a method to improve teaching and learning outcomes. These multiple studies serve to build on one another, providing a rigorous and increasingly comprehensive picture of when and how coaching works to impact educational outcomes. Notably, the studies discussed in this chapter include counterfactual quantitative studies, and more recent qualitative studies which attempt to tease out why and how coaching works, by examining the mechanisms at play.

While randomised control trials (RCTs) specifically have been considered the gold standard for measuring learning outcomes (Connolly et al 2018; Styles & Torgerson 2018), the use of RCTs in isolation has come under scrutiny more recently. As Deaton and Cartwright argue, “we cannot find out why something works by simply demonstrating that it does work, no matter how often, which leaves us uninformed as to whether the policy should be implemented” (2018, 61). The authors note that while RCTs provide valuable information, they should be just one piece of the puzzle. Further, Connolly et al. (2018) suggest that there are three serious criticisms of randomised-control-trial research in education. The first is that RCTs are blunt research designs that ignore or neglect context and experience. The second is that RCTs tend to generate simplistic universal laws of ‘cause and effect’. Finally, they are inherently descriptive and have made a limited contribution to theory building. The first criticism is linked to the problem of generalisability, that is, experiments are often conducted in non-typical circumstances often with participants who, by virtue of the fact that they are willing to consent to be part of a study, are unusual. Even if they are more or less typical of the target sub-population, extrapolating findings to a system population at large is problematic. The RCT studies are by definition designed to provide definitive information about the causal relationship between the intervention, treatment, or model and the desired learning outcome. More specifically, they are designed to provide information on whether the results are statistically significant or may have occurred by chance and the relative effect of the impact. Finally, given that many RCT studies had been conducted to ensure eligibility for government funding in the new wave of evidence-based policy, it is not surprising that some RCTs were initiated not to advance knowledge in education, but to access a ‘kosher’ stamp of approval from government funders.

In the last decade, researchers in South Africa have developed a long-term research programme anchored around large-scale counterfactual research studies that clearly and comprehensively addresses these three criticisms. The first major study of its kind was a Regression Discontinuity Design study undertaken on the Gauteng Primary Language and Literacy Intervention in 1,000 schools from 2011 to 2014. This was followed by the Reading Catch-up Programme (RCUP) study in 100 schools in 2014 which replicated one component of the Gauteng Primary Language and Mathematics Strategy (GPLMS), the Intermediate Phase large-scale remedial programme. The findings of this study fed into the Early Grade Reading Study I (EGRS I) (2015-2017), a multi-arm 240 school RCT that focused on early grade reading in the Home Language. The Early Grade Reading Study II (2017-2020) replicated the most effective change model, but focused on the teaching of reading in the First Additional Language in 180 schools. In the second intervention arm of the EGRS II, the research team tested the cost-effectiveness of virtual coaching compared to on-site coaching within a structured pedagogic programme. The final study, undertaken in 2018-2019, was undertaken by Funda Wande and included the basic change model but included a more intensive teacher-development component. Together the studies in this research programme are making a positive contribution to advancing knowledge in the field of early grade reading by showing the impact of a particular intervention model that is instructional coaching embedded in a structured pedagogic programme.

The multiple studies included in this chapter, when taken together, begin to present a full picture – not only of what works, but why it works (Deaton & Cartwright 2018). This study is providing knowledge that is not specific to any specific context but goes beyond simplistic conclusions that link interventions to specific learning outcomes, and contributes to theory building. Further, by presenting the findings over the past decade, this chapter also aims to flag questions that need to be addressed in the future.

The chapter begins with a brief review of the international literature on instructional coaching. It then examines the various initiatives and research studies that have incorporated instructional coaching into the change model. This is followed by an analysis of the sustainability of the results after the end of these initiatives. In section 2.6 we explore some of the qualitative research on embedded coaching, showing what has been learned about the mechanisms, which could contribute to theory building. The final section of the chapter signals potential areas that this research can address to debates in the wider field of education.

2 Literature review

There is a growing body of evidence to show, empirically, that coaching can work to change instructional practices and positively impact learner achievement, both in high-income contexts (Kraft, Blazar & Hogan 2017; Cohen et al. 2021) and low-income and middle-income contexts (Piper et al. 2018; Majerowicz & Montero 2018; Castro et al. 2021). In their widely cited meta-analysis of 60 empirically-driven studies, Kraft et al. (2017) found that, while coaching appears to have an overall significant impact on instructional practices (.49 SD) and a more moderate impact on student achievement (.18 SD), there is substantial variability in programme effect sizes. This variability is somewhat expected, given the wide practical variability in coaching interventions – and even what is defined as 'coaching'. The authors stipulate that the coaching process involves an "instructional expert work[ing] with teachers to discuss classroom practice in a way that is individualized..., intensive..., sustained..., context-specific... [and] focused" (Kraft et al. 2017, 9). However, even within this framework, coaching interventions can vary greatly in design. While the authors explore some of the design elements that may impact achievement, they do not offer a meaningful theory about why or how certain design elements may be more effective in impacting achievement than others. Critically, Kraft et al. (2017) show a negative correlation between programme size and programme effects, consistent with the well-established challenge of scaling up. However, again, descriptive details are needed to build meaningful theory about why this correlation exists with regards to coaching, and how it can be mitigated for more effective programme design at scale.

Another subsection of the coaching literature centres on principles for successful coaching models – the 'how to' literature. Knight (2007), for example, outlines his theory and seven key principles for a 'partnership approach' to instructional coaching. His principles, however, serve more as a theory of change, rather than a mechanism of how coaching actually works in practice. Further, while this 'how to' literature tends to be generic rather than context specific, it overwhelmingly comes from a Global North context. While the how-to coaching literature provides valuable insights to consider in intervention design and implementation (including coach training), it does not give insight into the processes at work in actual teacher-coach relationships. Similarly, situated in the broader literature reviewing structured pedagogic programmes is a theory of change that includes coaching without a deep understanding of how and why coaching works in practice. Coaching is discussed, essentially, as a 'catalyst' to ensure the enactment of the materials and resources are translated into new classroom practices (Fleisch et al. 2016; Fleisch 2016; Piper & Zuilkowski 2015; Piper & Zuilkowski 2014). In other words, there is a theory of how coaching leads to change, but not rigorous qualitative evidence of the mechanisms actually working to produce outcomes (or not).

Ultimately, while we have evidence that the inclusion of instructional coaching in educational change initiatives can significantly impact learning outcomes, far less is understood about how and why coaching works in practice. A growing body of research examines the dynamics of teacher-coach interactions, and the ways in which meaning is constructed in teacher-coach conversations. These studies pay careful attention to the ways in which coaches negotiate power and positioning (Rainville & Jones 2008), pressure and support (Ippolito 2010), and conversational content and moves (Heineke 2013) in their work with teachers. As Hunt (2016) notes, "literacy coaching involves complex emotional work within the physical and ideological spaces in which coaches and teachers interact" (Hunt 2016, 331). Hunt's (2016) work is part of a related but smaller body of research that examines the emotional spaces in which coaching interactions occur. These studies help us to understand the potential importance of emotionally supportive relationships for effective coaching work (Darby 2008; Vanderburg & Stephens 2010; Robertson et al. 2020). While these studies provide some suggestions of how coaching works, there is still work to be done in understanding the internal mechanisms that drive change – particularly in Global South contexts which are often associated with low-skill teachers and high class sizes. Research into the mechanisms of change in the Global South is critical, if we are to understand how coaching works in the context of large class sizes, resource shortages, underqualified teachers, and other contextualised challenges.

Globally, the results of coaching studies have shown some variable results. This is potentially rooted in a lack of consistency around the term 'coaching', and the design elements and specifics of coaching initiatives. It is important here to distinguish the coaches included in these research projects from subject advisors, who are sometimes trained by South African organisations to provide a supportive function. The coaches included in the research project below have content- and project-specific expertise, and are necessarily outside the formal education bureaucracy. Coaches' sole function is to support teachers to improve their teaching practice – they critically do not serve a monitoring or evaluative function. In terms of the research, this distinction is significant. In a South African context, instructional coaching embedded in structured pedagogic programmes has proven to be a key component of successful educational change initiatives. More recent qualitative studies are providing insight into how and why. This research is detailed below.

2.1 Early research on coaching as part of a structured pedagogic programme

The first research on instructional coaching in South Africa arose from the attempts to evaluate the Gauteng Primary Language and Mathematics Strategy (GPLMS) (2010-2014). One of the first of its kind in South Africa, the GPLMS was a government-led and -funded initiative aimed at improving learning outcomes in three key subjects, Home Language, English First Additional Language (EFAL) and mathematics in 1,000 no-fee primary schools. Selected schools were assigned to the intervention if their average Grade 3 English and mathematics scores fell below 40% on a provincial department standardised test conducted in 2008 (Fleisch, 2016). The intervention design made use of a 'best practice' for system-wide change with a combination of prescriptive lesson plans, the universal provision of quality learning materials, and both just-in-time training and one-on-one instructional coaching[1] for more than 8,000 teachers (Fleisch 2016). The intervention was implemented concurrently with the introduction of a universal testing system, referred to as the Annual National Assessments, as well as the introduction of the DBE Workbooks (see McKay & Spaull 2022) and a new curriculum – the Curriculum and Assessment Policy Statement (Muller & Hoadley 2019).

To assess the effectiveness of the intervention, researchers made use of a regression discontinuity design (RDD) approach, specifically comparing the change of relative school average score between comparable groups of selected schools. Primary schools with scores immediately above the 40% cut-off threshold were not assigned to the intervention and served as the 'control group', and schools that fell immediately below the cut-off 40% served as the 'treatment' group. While not measuring the impact of instructional coaching as an isolated component, the evaluation results showed that the model of instructional coaching embedded in a multi-component intervention model was very effective in improving learning outcomes at scale, at least for the types of schools around this 40% threshold (Fleisch & Schoer 2014; Fleisch et al. 2016).

Although not the immediate intention, the GPLMS intervention showed that this model of system-wide coaching as part of a structured pedagogic programme (Fleisch

1 Note that one NGO, Class Act, was directly involved in all four interventions but in different capacities. In the GPLMS it was one of 12 service providers contacted to manage coaches. In the RCUP and EGRS I they were the sole service provider including providing and managing coaches. In the EGRS II study they worked Molteno jointly to run the project.

2016, 2018) was financially viable. While a no cost-feasibility analysis (as defined by Levin & McEwan 2000) was conducted prior to implementation, the fact that the intervention was successfully conducted at scale and funded almost exclusively out of the official annual provincial budget strongly suggested that the model could be scalable in similar contexts.

Notwithstanding the important contribution that the GPLMS RDD study made, given the design limitations, conclusions could only be drawn about a local average treatment effect. The evidence from the design did not provide insights about the impact of the intervention model on schools at all points on the performance distribution, particularly schools at the bottom and the middle of the school performance distribution (Fleisch et al. 2016).

2.2 Reading catch-up programme

In a field not known for statistically significant effects for intervention models (Styles & Torgerson 2018), the GPLMS model itself and the strong effects shown in the RDD study created considerable interest in the sector. Given the limitation of the design of the evaluation, South African researchers began to explore more rigorous ways to evaluate the impact of the embedded instructional coaching model. The first such rigorous study was a replication study of a 'lite' version of the GPLMS model, a catch-up programme that used the embedded model of instructional coaching with lesson plans and quality materials, focusing on English as a First Additional Language in Grade 4. The catch-up programme that had been implemented in Gauteng in 2012 had exhibited large average learning gains between baseline and endline testing. Using the same structured lesson plans, storybooks, coaching, and service provider as in GPLMS (i.e. Class Act), and in the same dosage, the research team designed a randomised control trial to rigorously evaluation the impact of the catch-up programme in a single district of Kwazulu-Natal (Fleisch 2017). This research team included both university and government researchers.

Unlike the Gauteng intervention, the results of the Kwazulu-Natal RCT showed no statistically significant learning or educationally meaningful difference between the control and intervention school groups at the endpoint of the intervention programme.

When average scores were disaggregated for subgroups, it emerged that the intervention schools with lower baseline scores improved less than comparable control schools.

The conclusion drawn from the subgroup analysis was that, while the intervention may have been appropriate for schools with stronger average baseline learning outcomes, the programme was not at the right level for learners with weaker learning outcomes in weaker schools at the outset of the study. In other words, the intervention model with embedded coaching shows promise, but to benefit all, careful consideration needs to be given to ensure that the intervention is aligned to the average learning profiles or pitched at the right learning level, and targeting the right language. This is an insight that has become increasingly important in the past few years in the international literature (see Kaffenberger 2019 for example)[2].

2 Although the study did not give us critical information about learning inequality, it is possible that the top 40% of the distribution had acquired some basic reading skills in HL and could benefit from an EFAL intervention while the weakest 60% had not learned the requisite literacy skills in Grades 1-3 and might consequently fail EFAL irrespective of the intervention.

From the perspective of instructional coaching, the results pointed to the need to look closely not just at coaching as a model but specifically at the characteristics and practices of individual coaches. Controlling for baseline scores, the data showed that the two coaches in the programme, who used the same structured programme, had a different level of impact on the schools they worked in. This evidence suggests that the impact of an intervention that uses coaches to support teachers may depend on the ability and experience of staff doing the coaching. In other words, it is likely not just the presence of a coach, but the actual content of coaching that matters. Further, this discrepancy points to the need for qualitative studies to accompany rigorous quantitative measures – to give a deeper and more complete picture of the achievement outcomes.

2.3 Early Grade Reading Study I

Adopting the GPLMS implementation model that included instructional coaching, the first Early Grade Reading Study, the largest and most ambitious randomised control trial at that time in South Africa, led by the Department of Basic Education, was designed to test the basic change model – instructional coaching combined with daily lesson plans and quality educational materials. The study allowed for comparative analysis of three distinct intervention models: the basic model with lesson plans, educational materials, and training that included embedded coaching; a second model that had all the components of the basic model other than instructional coaching; and a completely different approach, which worked directly with parents to help them support reading at home (Cilliers et al. 2020).

What, exactly, did instructional coaches provide as part of the basic model of the intervention? In this model, trained reading coaches visited each teacher on a monthly basis over the duration of the academic year. The purpose of the coaching was to support teachers to implement the core methodologies in the daily lesson plans provided, thereby improving their content knowledge, pedagogical techniques, and ultimately, helping to build teachers' professional confidence. During the classroom visits, the coaches would typically observe teaching, provide feedback on how to improve the lesson, and demonstrate or model correct teaching techniques. The coaches often held small-group sessions at the start of each of the term to orient the teachers to the new learning materials and lesson plans for that term. They also occasionally held afternoon workshops for teachers at clusters of schools. While the GPLMS intervention employed more than 500 coaches, the EGRS I intervention employed three coaches, each serving 16-17 schools. The coaches all had extensive experience in Foundation Phase classrooms, had a minimum of a bachelor's degree, and received regular training and ongoing support from a coach supervisor.

After two years of intervention in Setswana home-language reading in 50 schools, the basic model with instructional coaching embedded in a structured pedagogic programme showed the strongest and most consistent gains overall, and in all but one of the sub-test outcomes. The total study included 230 schools (80 control, 50 each of the three intervention models). Researchers estimated that the impact within the two-year intervention was equivalent to approximately a third to half a year of schooling.

In a recently published follow-up study conducted with the same teachers and original cohort of learners, Cilliers et al (2022a) confirmed that the learning gains made in the early years of the EGRS I study persisted. One year after participating in the programmes, teachers in both of the structured pedagogic programmes (coaches and LTSM on the one hand and centralised training and LTSM on the other) retained knowledge and continued using the learning resources they had been provided as part of the EGRS I intervention. However, only teachers who had participated in the coaching model continued to use the improved teaching techniques in their classrooms. The magnitude of the effects on learners taught by coached teachers one year *after* the programme ended was 0.19 SD, roughly half the size relative to those at the end of Grade 2. These significant 'fade-out' effects, even after one year, add weight to arguments for the need for sustained coaching interventions for sustained impacts.

The EGRS I study provided the first conclusive proof that instructional coaching 'works' to improve early grade reading in the home language and that without the coaching component, intervention models were unlikely to have much impact.

2.4 Early Grade Reading Study II

EGRS I, conducted in the North West province, was a research success. It clearly demonstrated the usefulness of RCTs, it showed that one particular change model – embedded coaching within a structured pedagogic programme – could lead to substantial improvement in early grade reading outcomes in children's home language, and showed further that the intervention effects would not fade out quickly. But two key questions remained. First, would the results that first emerged in the RDD study in Gauteng, and then in the EGRS I in the North West, be replicable? That is, would the same basic model have the same learning effects in other locations or provinces and in other subjects? Second, given the relatively high cost of embedded coaching, could an alternative form of coaching, particularly a form of coaching that was lower cost and easier to implement in relatively sparsely populated areas be equally effective?

It was these two questions that motivated the second Early Grade Reading Study (EGRS II). Using the same basic change model that incorporated on-site coaching, daily lesson plans and quality materials, and a second model that used virtual or e-coaching[3], lesson plans and quality materials, the study set out to investigate the learning impact of these models on learning outcomes in early grade learning of English as a first additional language (EFAL) in the province of Mpumalanga. As with the EGRS I, the study used a sample size of 50 schools each for the intervention arms (i.e. the on-site coaching arm and the virtual coaching arm) and 80 schools for the control group. Beginning with a pilot study in 2015, the full implementation took place from 2016 to 2019, and followed one cohort of learners through Grades 1 to 3.

The in-person coaching model was consistent with the embedded coaching in the EGRS I. Teachers in the virtual coaching model received a tablet with an electronic version of the lesson plans, and received the same learner reading materials and other classroom resources at face-to-face training twice per year. The virtual coach called

3 In the virtual coaching intervention, one virtual coach supported about 80 teachers from afar. The virtual coach called each teacher at least once during a two week cycle – with teachers classified as struggling receiving additional phone calls. In addition, the virtual coach ran a competition via WhatsApp, in which teachers were to submit photographs or videos of their teaching, every two weeks.

teachers on a regular basis and sent weekly reminders and teaching tips through WhatsApp. The virtual coach called every teacher at the start of the term, and followed up every few weeks, with more intensive interaction if she felt that the teacher required additional support. In addition to the lesson plan, the tablets included additional electronic resources such as short videos of the teaching methodologies, sound clips of the phonics sounds, songs and rhymes, and examples of learners' work. The virtual coach also introduced small competitions around specific themes via WhatsApp groups. Teachers were required to submit either video clips of their teaching or photographs of learners' work for the competitions. The coach would then choose the best teacher in each of the teacher groups, who was then awarded with a small amount of airtime. The competitions were intended to give the virtual coach a way to observe actual teaching practice, thus enabling her to provide more targeted feedback. The competitions also helped teachers to see what other teachers in similar contexts were doing, thereby fulfilling the role of a virtual community of practice.

After the first year of schooling, it appeared as if the on-site and virtual coaching would have roughly similar and strong learning gain outcomes compared to learners in the control schools (Kotze et al. 2019). However, the pattern of these results shifted dramatically by the end of Grade 3 (Cilliers et al. 2020b). Learners in the on-site coaching group made strong gains in oral proficiency in English and statistically significant, albeit more modest gains in English reading by the end of Grade 3. In the reading subtests, the learners in the coaching schools made significant gains in oral reading comprehension, a key indicator of reading for meaning. However, children in the virtual coaching schools appeared to perform more or less like children in the control schools, and may even have performed worse than these children in home language (not a focus of the intervention), suggesting that virtual coaching may have had a negative spill-over.

As with the EGRS I, follow-up research showed that the learners in the basic model intervention continued to show sustained improvement long after the implementation team had completed their work in Mpumalanga. Notwithstanding fadeout, children in the coaching group showed higher levels of core comprehension, listening, and speaking English comprehension skills, with no evidence of any impact in the virtual coaching intervention schools at the end of Grade 4 (DBE, 2022). Given that the follow-up study was undertaken during Covid, the scale of the fadeout should be interpreted with caution.

2.5 Funda Wande (Eastern Cape)

Much of the research on instructional coaching as part of an integrated structured pedagogy programme has been done under the banner of either the GPLMS or EGRS. However, there is now a separate line of research in South Africa on the impact of instructional coaching. Funda Wande, an NGO, assessed the impact of a structured pedagogic programme in the Eastern Cape centred on a far more intensive in-service training programme (a more intensive coaching model) that included materials and structured guidance for teaching. The training programme included extensive video tutorials and enrolled the Heads of Department (HODs) in a two-year university-certified programme focusing on teaching literacy in the Foundation Phase (see Murray & Taylor 2022). Funda Wande's more intensive coaching model comprised six expert coaches who were experienced foundation-phase literacy educators, with a coach to school ratio of 1:5. The coaches observed Grade 1-3 teachers three times a month in their classrooms, and provided teachers with targeted advice on how to improve their practice, as well as providing model lessons with their learners.

Due to Covid-19, the midline results of this RCT are probably not the most accurate reflection of what the intervention is likely to achieve in normal circumstances. The limitation is that the intervention was only active for a single year and measured only the gains in Grade 1 and 2. Ardington and Meiring (2020) concluded that the Funda Wande intervention had a statistically significant 0.17 SD impact on the learners' reading proficiency after the first year and the positive overall effect was driven by an improvement in all the sub-domains of reading proficiency.

2.6 Qualitative case studies on coaching and virtual coaching

In a seminal paper on RCTs, Deaton and Cartwright (2018) argued that:

> *Without the structure that allows us to place RCT results in context, or to understand the mechanisms behind those results, not only can we not transport whether 'it works' elsewhere, but we cannot do one of the standard tasks of economics, which is to say whether the intervention is actually welfare improving. Without knowing* **why** *things happen and* **why** *people do things, we run the risk of worthless casual ('fairy story') theorizing and have given up on one of the central tasks of economics and other social sciences*
>
> (2006, 18).

RCTs have been criticised for the absence of a concern about advancing theory, or uncovering multiple causal mechanisms. Critics of RCTs have consistently argued that the value-add of this method of enquiry is limited. To address this limitation, the early grade learning programme has begun to incorporate qualitative case studies to probe for the multiple mechanisms that would account for the cost-effectiveness of embedded coaching in multi-component structured pedagogic models.

Fleisch and Dixon's (2019) case study of mechanism in the Early Grade Reading Study did not explicitly interrogate the role of on-site coaching but rather it explored possible mechanisms in the model of change as a whole. Interviewing teachers, school managers and observing classrooms, they paid close attention to the impact on teachers who used the scripted lessons. They observed that adopting the methodologies in the lesson plans led to new routines and rhythms directed on teachers' and learners' bodies. They noted that EGRS I teachers mostly kept to lesson-plan time and learners rapidly became habituated to sequences of activities that followed on each other and they became familiar with the set of tasks associated with the string of activities in the lesson plans.

For example, the researchers saw the ways designated groups of learners purposefully moved to the carpet for group guided reading. The adoption of prescribed and timed activities and cycles of repetition resulted in increased instructional time within the lesson. The step-by-step timed nature of the lesson plans was enabling for teachers. Breaking tasks down into timed steps has the benefit of breaking time into manageable pieces that work to increase overall academically focused time. When learners and teachers internalise routines, and transitions are smooth, several operations can take place simultaneously. These insights were confirmed when the researchers compared the use of time during the Setswana lessons to time wasted during the numeracy and EFAL lessons. This insight was confirmed in the second set of case studies conducted in the Early Grade Reading Study II (Botha and Schollar 2018). In that report, the issues about macro-time were highlighted in relation to the use of the lesson plans.

While the use of time and quality of instruction generally improved, all cases studies showed that take-up was uneven across classrooms. Efficient new use of time was affected by older practices, relative comfort with new practices, and some of the design flaws in the CAPS curriculum and the EGRS programme. While the teachers followed the lesson plan steps, it was clear that they did not all have real mastery of instructional tasks. The researchers noted that eight months into the year teachers still relied on lesson plans and referred to them often.

While time was a key factor that linked the programme to improved learning, Reeves' (2017) case study observations and interviews at the end of Year 1 in the EGRS II pointed in a very different direction to explain possible mechanisms associated with different models of instructional coaching. Reeves observed real differences between practices she saw in the control schools and the two intervention-group schools with coaching. Specifically, she noted that lessons in virtual coaching classes could have been substantially improved had the teacher had access to an on-site coach in the classroom. Of particular interest is her finding that the direct school visits by the EGRS II on-site coaches and their discussions with principals or SMTs changed the attitudes of the entire Foundation Phase staff and the school management teams.

Arguably the most powerful insights into mechanisms emerged from Alsofrom's doctoral study. Alsofrom (2019) explores the critical question: how and why do structured pedagogic programmes work to change teachers' practices? Through a case-study approach using classroom observations and in-depth interviews, her study examines the mechanisms through which the EGRS II multiple components work to impact change. In her study, she develops a framework for looking at the mechanisms through which the EGRS impacted teachers' practices through a non-linear, cyclical process, occurring in three stages[4]: reaching emotional readiness (stage one); creating a concrete and meaningful language of teaching and learning (stage two); witnessing change (stage three). In stage one, teachers expressed a range of emotions including fear, anger, humiliation, embarrassment, shame, and incompetence. She found that on-site coaching serves to recognise these emotions and provides an emotional space that both acknowledges the feelings and mitigates them through consistent and appropriate positive feedback by affirming teachers' identities with encouragement and empathy. This new emotional space allows teachers to be vulnerable, and in so doing, allows them to learn a new practice.

In the second stage, she found that change was driven by supportive teacher-coach

4 This framework developed in Alsofrom's work was developed retrospectively and emerged from her indepth qualitative research. As such the programmes were not designed with the framework in mind.

relationships, which created the necessary conditions for meaningful learning around discussions about 'mistakes' and taking risks. In stage two, implementing educational interventions involved a complex process of learning – not simply mechanical implementation. In this stage, the coaching process allows teachers to concretely see and experience the new teaching approaches. As teachers developed a tangible and meaningful understanding of key literacy methodologies, their practice began to change. For example, they were able to translate a technical term like phonics or group guided reading into a set of concrete activities in their own classrooms.

In the third stage, as the teachers became more confident in the use of the new and more effective practices, they began to observe the impact of the practices on their own learners in the classroom. This was a critically important positive intrinsic incentive, which the coach attributed both to her own hard work and improvement – and to the EGRS II programme – for helping support her growth and development. While Alsofrom found the teachers were increasingly confident and competent at this stage, the teachers receiving on-site coaching were still reliant on the coach for reassurance. In her view, during the third stage of coaching, the coaches increasingly played the role of agents of professional accountability – which teachers in this stage sought out. As Volmink (DBE 2016) notes in his report, accountability is weak in the South African education system. The accountability that coaching introduces into the South African system is tied to the desire to impress – and not to disappoint – someone (the coach) who takes a keen interest in your work and change over time. Again, this type of accountability is rooted in the emotional connections and trusting relationships that are critically built in stage one.

Alsofrom's observations align with the Botha and Schollar (2018) case-study research on both on-site and virtual coaching. Along with a range of insights, they conclude:

> *What might be an argument against virtual coaches is based on one of its greatest apparent strengths – the ability of teachers to request help and advice on any topic without having to wait for a school visit. This certainly does increase the immediacy of available support but it relies ultimately on the teachers' knowledge of their own shortcomings. As the popular phrase has it, how do you know what you don't know? On-site coaches, on the other hand, observe how teachers actually present lessons in classrooms and are ideally placed to recognise problems and solutions to teacher activity. This ability to observe teacher 'dysfunctions' or misunderstandings in action is lost to the virtual coach who instead must rely on self-reported problems.*
>
> (Botha & Schollar 2018, 17)

Botha and Schollar's observation points to one of the central strengths of the on-site coaching model, the ability to directly observe practice, and ascertain first-hand what problems and challenges the teacher faces in her practice. The case studies of the EGRS II showed that although information technology has the advantage of immediacy, it does not provide the space for direct scrutiny and live support in the classroom.

In summary, these qualitative case studies uncover a complex web of mechanisms that sit within the embedded instructional coaching within a structured pedagogic programme. Some are best thought of as enabling factors such as managing micro- and macro-time, which makes instructional impact more likely. Others, such as facilitating a community of practice, are much better classed as a mechanism. The work on the role that coaching plays in setting the emotional foundation for professional learning in situ is a new and important contribution to understanding the specific mechanisms associated with coaching. Coaching also provides practice-based teaching and learning in a way that enables teachers to take the new practices forward.

3 Conclusion

What can be concluded from the review of the early grade reading research on coaching in South Africa? Firstly and most importantly, the research programme has incorporated multiple quantitative and qualitative studies of an intervention model that was largely consistent in design, implementation fidelity, and dosage with very similar findings across interventions, representing a major advancement in the knowledge about system-wide educational change in South Africa. This claim is further supported by the inclusion of in-depth case studies that provide compelling explanations of causal mechanisms associated with this intervention model. The two major criticisms of RCT research are that it is strong on internal validity but weak on external validity, and that it provides no insight into the mechanisms of change. These criticisms have been taken seriously and addressed systematically. The critiques of RCT research are one of the most important assets that the researchers have mobilised to enhance this research programme.

The second, and possibly one of the strengths of this research programme, has been the uncovering of multiple overlapping and mutually re-enforcing change mechanisms introduced with instructional coaching. Within the development-economics literature, there is a tendency to conflate the intervention itself with the mechanism associated with the intervention and privilege a single mechanism. The early grade reading research, particularly the research that incorporates the study of instructional coaching, shows very clearly not only that the intervention model works, but it also suggests that multiple mechanisms work together to help improve learning outcomes.

Thirdly, the research programme draws on and speaks back to the wider literature (theory and practice) on educational change. This overcomes one of the criticisms that Shalem and De Clerqc (2022) have levelled against RCT research.

The obvious starting point for locating the early grade reading research in the wider literature is to link it to the debates about alignment and coherence as a key factor in education improvement. In many respects, this debate was initiated in the United States by Fuhrman (1993), and taken up by Cohen et al (2003). The key idea here is that change requires aligned and coherent 'instructional guidance'. Without it, substantial system improvement in instruction would be near to impossible because of conflict and confusing signals from policy. The early grade research presented here not only signals the importance of alignment and coherence but goes one step further to suggest which of the various components (specifically coaching) within an aligned intervention may be central and which may be peripheral. In this case, the research provides a fairly strong knowledge claim for the centrality of on-site in-classroom instructional coaching.

Another example is one of the central theories associated with the mechanisms of lesson plans, which is offered by Hiebert and Morris (2012), who locate the lesson-plan debate within the wider context of strategies to improve classroom instruction. They suggest that lesson plans could and should focus on teacher knowledge.

In contrast, the findings from the early grade reading research suggests that in practice the primary mechanism of lesson plans may be different; that is, while some teachers may increase their professional knowledge and know-how when using lesson plans, the key mechanisms associated might be related to the improved use of micro- and macro-time, i.e. helping teachers structure their daily, weekly, and termly teaching schedules.

Similarly, the research on the difference between on-site and virtual coaching has revealed the centrality of emotions and building trust in the change journey. The on-site coaching relationship is about more than just the one-on-one contact between the teachers and the coach, and more than just instructional conversations – it involves creating emotionally supportive, trusting relationships, and strengthening social networks around curriculum and pedagogic conversations between teachers and other teachers, and also with school managers. This insight needs to be located in relation to the wider literature on educational change and social network theory (Daly et al. 2010; Datnow 2012).

When we deepen our interrogation of the structured pedagogic programme and its links to instructional coaching, another important feature of change emerges. That feature is accountability. While Barber (2005) makes the case for external accountability mechanisms and Darling-Hammond (1989) for professional accountability standards, a major shift in the debate about accountability has focused on internal accountability. Fullan et al. (2015) have suggested that the most powerful form of accountability is what they refer to as professional capital, i.e. the deep knowledge, skill, experience and expertise that educators need to acquire and then apply when working with colleagues. The coaching process in the South African early grade reading studies, while not necessarily achieving the high level of professional capital advocated by Fullan et al. does point towards the centrality of professional capital for meaningful change.

References

Alsofrom, K. (2019). How and why does coaching work to improve teaching practices in the EGRS2? An examination of causal mechanisms. UKFIET conference.

Ardington, K. and Meiring, T. (2020). Midline Findings of the Impact Evaluation of the Funda Wande Coaching Intervention. Cape Town: SALDRU, University of Cape Town. Available at: https://fundawande.org/news/impact-evaluation-of-funda-wande-coaching-intervention-midlinefindings-23 (accessed 7 July 2020).

Barber, M. (2005). The virtue of accountability: System redesign, inspection, and incentives in the era of informed professionalism. *Journal of education, 185*(1), 7–38.

Botha, D. & Schollar, E. (2018). Case Studies in EGRS II Schools 2018. Early Grade Reading Study. Available at: https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/837 (Accessed 14 November 2022).

Castro, J. F., Glewwe, P., Heredia-Mayo, A. & Montero, R. (2021). *Work with What You've Got: Improving Teachers' Pedagogical Skills at Scale in Rural Peru* (No. 1701-2021-3436).

Cilliers, J., Fleisch, B., Kotze, J., Mohohlwane, M., & Taylor, S. (2022). The challenge of sustaining effective teaching: Spillovers, fade-out, and the cost-effectiveness of teacher development programs. *Economics of Education Review. 87*, 102215.

Cilliers, J., Fleisch, B., Kotze, J., Mohohlwane, N., Taylor, S., & Thulare, T. (2022a). Can virtual replace in-person coaching? Experimental evidence on teacher professional development and student learning. *Journal of Development Economics*. 102815.

Cilliers, J., Fleisch, B., Prinsloo, C., & Taylor, S. (2020). How to improve teaching practice? An experimental comparison of centralized training and in-classroom coaching. *Journal of Human Resources*. 55(3), 926–962.

Cohen, D. K., Raudenbush, S. W., & Ball, D. L. (2003). Resources, instruction, and research. *Educational evaluation and policy analysis. 25*(2), 119–142.

Cohen, J., Krishnamachari, A., & Wong, V. C. (2021). Experimental Evidence on the Robustness of Coaching Supports in Teacher Education. (EdWorkingPaper: 21-468). Available at: https://doi.org/10.26300/dgf9-ca95.

Connolly, P., Keenan, C., & Urbanska, K. (2018). The trials of evidence-based practice in education: A systematic review of randomised controlled trials in education research 1980–2016. *Educational Research*. 60(3), 276–291.

Daly, A. J., Moolenaar, N. M., Bolivar, J. M., & Burke, P. (2010). Relationships in reform: The role of teachers' social networks. *Journal of educational administration*. 48 (3), 359–391.

Darby, A. (2008). Teachers' emotions in the reconstruction of professional self-understanding. *Teaching and Teacher Education, 24*(5), 1160-1172.

Darling-Hammond, L. (1989). Accountability for professional practice. *Teachers College Record, 91*(1), 59–80.

Datnow, A. (2012). Teacher agency in educational reform: Lessons from social networks research. *American Journal of Education, 119*(1), 193–201.

Deaton, A., & Cartwright, N. (2018). Understanding and misunderstanding randomized controlled trials. *Social Science & Medicine. 210*, 2–21.

Department of Basic Education (2016). Report of the Ministerial Task Team on Investigation into the Allegations into the Selling of Posts of Educators by Members of the Teachers Unions and Department Officials.

Department of Basic Education (2022). The Second Early Grade Reading Study Year 4 Draft Report Evidence one year after implementation

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 Malone, H.J., Rincón-Gallardo, S., Kew, K. (Eds), *Future Directions of Educational Change*. New York, NY: Routledge.

Fleisch, B. (2018). T*he Education Triple Cocktail: System-wide Instructional Reform in South Africa*. Cape Town: UCT Press/Juta and Company (Pty) Ltd.

Fleisch, B., & Dixon, K. (2019). Identifying mechanisms of change in the early grade reading study in South Africa. *South African Journal of Education,* 39(3).

Fleisch, B., & Schoer, 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), 1–12.

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

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

Fullan, M., Rincón-Gallardo, S., & Hargreaves, A. (2015). Professional capital as accountability. *Education policy analysis archives, 23*(15), 1–21.

Heineke, S. F. (2013). Coaching discourse: Supporting teachers' professional learning. *The Elementary School Journal*, 113(3), 409–433.

Hiebert, J., & Morris, A. K. (2012). Teaching, rather than teachers, as a path toward improving classroom instruction. *Journal of teacher Education, 63*(2), 92–102.

Hunt, C. S. (2016). Getting to the heart of the matter: Discursive negotiations of emotions within literacy coaching interactions. *Teaching and Teacher Education, 60*, 331–343.

Ippolito, J. (2010). Three ways that literacy coaches balance responsive and directive relationships with teachers. *The Elementary School Journal, 111*(1), 164–190.

Kaffenberger, M. (2019). A Typology of Learning Profiles: Tools for Analysing the Dynamics of Learning. *RISE Insight Series, 15*.

Knight, J. (2007). *Instructional coaching: A partnership approach to improving instruction*. Thousand Oaks, CA: Corwin Press.

Kotze, J., Fleisch, B., & Taylor, S. (2019). Alternative forms of early grade instructional coaching: Emerging evidence from field experiments in South Africa. *International Journal of Educational Development*. 66, 203–213.

Kraft, M. A., Blazar, D., & Hogan, D. (2017). The effect of teacher coaching on instruction and achievement: A meta-analysis of the causal evidence. *Review of educational research*, 88(4), 547–588.

Levin, H. M., & McEwan, P. J. (2000). *Cost-effectiveness analysis: Methods and applications* (Vol. 4). Thosand Oaks, CA: Sage.

Majerowicz, S., & Montero, R. (2018). Can Teaching be Taught? Experimental Evidence from a Teacher Coaching Program in Peru. *Job Market Paper*.

Muller, J., & Hoadley, U. (2019). Curriculum reform and learner performance: An obstinate paradox in the quest for equality. In Spaull, N. & Jansen, J. (Eds), *South African schooling: The enigma of inequality*. New York, NY: Springer.

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

Piper, B., Zuilkowski, S. S., Dubeck, M., Jepkemei, E., & King, S. J. (2018). Identifying the essential ingredients to literacy and numeracy improvement: Teacher professional development and coaching, student textbooks, and structured teachers' guides. *World Development*, *106*, 324–336.

Rainville, K. N., & Jones, S. (2008). Situated Identities: Power and positioning in the work of a literacy coach. *The Reading Teacher, 61*(6), 440–448.

Reeves, C. (2017) Grade 1 English First Additional Language Study in Mpumalanga Province 2017. Available at: https://www.education.gov.za/Portals/0/Documents/Publications/EGRS/EGRS%20II%20Website%20Upload/Reports/3_EGRS%20II%20EFAL%20Case%20Study%202017_Final_WEB.pdf (accessed 14 November 2014).

Robertson, D. A., Padesky, L. B., Ford-Connors, E., & Paratore, J. R. (2020). What Does It Mean to Say Coaching Is Relational? *Journal of Literacy Research*, 52(1), 55–78.

Shalem, Y. and De Clercq, F. (2022 in press) Randomized control trials in education (RCTs): What is their contribution to education theory about teaching? *Journal of Education.*

Styles, B., & Torgerson, C. (2018). Randomised controlled trials (RCTs) in education research–methodological debates, questions, challenges. *Educational Research*, 60(3), 255–264.

Vanderburg, M., & Stephens, D. (2010). The impact of literacy coaches: What teachers value and how teachers change. *The Elementary School Journal, 111*(1), 141–163.

04

Professional development to strengthen early grade reading: Lessons from the Primary School Reading Improvement Programme (PSRIP)

DHIANARAJ CHETTY, PENELOPE GROOME, KAITLIN ALSOFROM & JOHN THURLOW

Abstract

This chapter reviews the programme design and content of the Primary School Reading Improvement Programme (PSRIP), a national-scale reading intervention initiated by the Department of Basic Education (DBE) in 2016 and implemented through the National Education Collaboration Trust (NECT). It also looks at the evidence generated by the programme on teaching reading in English First Additional Language (EFAL) professional development outcomes.

Part I of the chapter explores the questions of what programme design choices were made and why, and how these have evolved over the years. Furthermore, it critically examines the implementation issues arising from working in collaboration with provincial government as well as with district officials and teachers. Part II of the chapter examines the available data on the pre- and post-assessment results of training subject advisors and educators who participated in the programme. The paper concludes by analysing the implications for teacher professional development in teaching reading.

KEYWORDS
EFAL, reading programme, PSRIP, NECT, professional development

1 Introduction

Teacher professional development is a critical component of reading improvement programmes. Whilst the evidence base on reading improvement interventions in South Africa has expanded rapidly in recent years, there are gaps in our understanding of the design issues in teacher professional development. Furthermore, the vexed issues are often found in how interventions translate into programmes that are implemented in ‘real world’ conditions – where teacher development has to be delivered at scale in widely differing contexts that challenge programme fidelity.

The PSRIP was one of a range of national responses to the outcomes of the PIRLS 2016 report in which South Africa’s low reading achievement levels were highlighted. Adopted as a sector plan by the DBE, the PSRIP was designed to strengthen teaching reading in EFAL through a large-scale teacher-development initiative, using learning programmes (including lesson plans, trackers, and supporting materials), which incorporated the latest instructional methodologies and materials based on evidence from other research studies in teaching reading.

Part I of this chapter explores the questions of what programme-design choices were made and why, and how these evolved over the years. Furthermore, it critically examines the implementation issues arising from working in collaboration with provincial government as well as with district officials and teachers.

Part II of this chapter examines the available data on the pre- and post-assessment results of training subject advisors and educators who participated in the programme. These are short-term outcomes which provide insight into the profiles of participants. Furthermore, they simulate the data on which provinces typically rely in gauging the outcomes of teacher professional development programmes. The chapter concludes by analysing the implications for teacher professional development in the teaching of reading.

2 Literature review

The literature on the design and outcomes of early grade reading programmes in South Africa has expanded rapidly in the past 20 years, much of it driven recently by the government’s own Early Grade Reading Study (EGRS) which was conducted at scale. Other smaller-scale interventions have also reported results which have contributed to the evidence base and to vigorous debate (Funda Wande, 2020). With a stronger focus on programme design and the pedagogy of teaching reading, Fleisch (2018) has made the strongest case for structured learning programmes which are at the core of global evidence (Chakera et al. 2020). The 2030 Reading Panel has made wider system-level recommendations based on an assessment of progress in the sector (2030 Reading Panel 2022). Overall, the pendulum has swung towards measurements of impacts in contrast to programme design in recent research.

3 Part I

PSRIP: Design and Training Review

At its inception, the PSRIP had the benefit of building on the experience of a number of large-scale efforts at improving reading in the early grades, including Foundations for Learning Campaign 2008-2011 (DBE); the Gauteng Primary Language and Mathematics Strategy (GPLMS) and the Programme for Improvement of Learning Outcomes (PILO). While much had been learned over the years in terms of programme content and pedagogy, many of the challenges in programme design remained. Questions such as the following remained unanswered: how to design for scale in a system as large and diverse as South Africa's? How to make effective use of human capacity already in the system? How could the programme successfully change practices in the short term and institutionalise itself over the longer term? Importantly, the PSRIP was required to respond to the needs of subject advisors and Foundation Phase specialists in the system on 'what' and 'how' in programme design. This would be a journey that continued to evolve in successive iterations of the programme until 2020. To start with, this paper focuses on content, pedagogy, and instructional design issues.

The design team had, by the start of the project, reviewed a wide scope of literature on education interventions across the Global South. By this time, structured pedagogical interventions had already emerged as a significant and successful category of education interventions (Conn 2014; Evans & Popova 2016; Snilstveit et al. 2016; Piper et al. 2015) and the design team had similarly used a structured pedagogical approach in two previous large-scale interventions (GPLMS, 2010 – 2013; EGRS I 2015 – 2017).

As Kim and Davidson (2019) note, structured pedagogy has been proven to positively impact learning outcomes because 'it offers learners evidence-based, effective learning opportunities to practise and acquire core skills' (2019, 2) – and when the principles of structured pedagogy are applied to reading instruction, learners have 'ample opportunities to become experts in essential tasks' – ultimately becoming skilled readers. From the outset of the PSRIP, the DBE and NECT decided that the Structured Learning Programmes (SLPs) designed as part of the structured pedagogical approach would be working documents that would evolve over time, in line with emerging evidence of best practices in the literature, and as teachers' knowledge and skillsets evolved.

While some of the content and design features of the SLPs in the PSRIP have changed over time – in response to participant feedback, review processes, budgeting realities, and additional research done by the design team – the original, basic structure of the SLPs has essentially remained unchanged. At the core of the SLPs are strong routines and core methodologies, which aim to change the way teachers use time in their classrooms, and to help teachers habituate new, improved practices through structure and repetition.

As Fleisch and Dixon (2019) note in their study of the EGRS I, the structured routine and new lesson plans helped shift classroom practices by transforming teachers' use of time and space in the classroom. Studies across the Global South, including South Africa, have shown that slow curriculum pacing (Taylor 2019) and an outdated and ineffective set of pedagogical practices such as drilling, choral recitation and copying from the board are pervasive in South African classrooms (Fleisch 2016; Piper & Zuilkowski 2015).

At a basic level, new (structured, fast-paced) routines and core methodologies aim to provide manageable and user-friendly solutions for even the most under-trained teachers, ultimately seeking to break the reliance on rote-learning methodologies in teaching reading. In addition, the programme was designed to explicitly foster social and emotional engagement and learning at all levels – between trainers and beneficiaries, between teachers and learners, and between learners and learners.

At a macro-level, a set of core methodologies in the SLPs, which are repeated with new content each week, was designed to ease the adoption of new teaching practices for teachers. Such repetition allows a teacher to practise a methodology again and again, ultimately reaching mastery in implementation. While implementation of these new methodologies may be shallow at the outset, the repetition means that, in the long term, these strategies can lead to deeper execution and understanding as implementation of the practices improves and theoretical knowledge is scaffolded through training and support. Then, at the micro-level, each of the core methodologies in the PSRIP programme was designed to align with evidence-based best practices for teaching reading, within the constraints of:

- budgetary realities
- the CAPS curriculum
- resource-limited classrooms
- existing teacher knowledge and skill sets.

In addition, because the SLPs were designed with large-scale, system-wide change in mind, the design of each methodology was developed with the aim of being workable for even the most under-trained and / or inexperienced teacher.

As an example of the design process, the Shared Reading methodologies went through several iterations, based on additional research of evidence based best practices for teaching reading. After an analysis of the CAPS curriculum and the evidence base on teaching reading (for example: Dehaene 2009; National Reading Panel 2000; Kilpatrick 2015; The Reading League 2022), the design team made the decision to designate Shared Reading as a time for building the oral language and comprehension skills that have been shown to be critical for learners' reading comprehension – specifically, building learners' vocabulary and demonstrating the internal thought processes of fluent readers – and decidedly not as a time to focus on learners' own technical reading skills.

It is critical to note here that, while the Shared Reading guidelines in CAPS suggest practices like using a 'simple enlarged text' so that learners may follow along with the words as the teacher reads the text aloud and points, the scientific research on reading has determined that such practices are not, in fact, useful to learners' own word-reading skills (Kim 2020; Evans & Saint-Aubin 2005). PSRIP therefore limited the scope of Shared Reading to the development of reading comprehension skills and not to the development of technical decoding.

Further, it was decided that Shared Reading be used as a critical time to expose learners to rich texts with complex storylines and to build joy and love in the experience of reading. While this would be important in any classroom, the urgency of such exposure was deemed to be high given the scarcity of rich narrative texts both in classrooms and in many learners' homes (Tella & Akande 2007; Mugambi 2015; Pretorius & Naudé 2002). Based on both CAPS requirements (for example, the requirement to talk with learners about the pictures in the story) and the research on repeated and interactive read-aloud strategies (Gold & Gibson 2001; McGee & Schickedanz 2007), Shared Reading was designed to include a pre-read, first-read, second-read, and post-reading activity. This repeated exposure to the story is designed based on the research-based approach to using reading aloud to build deeper comprehension and vocabulary in children's early schooling – particularly children with limited exposure to literacy experiences at home (McGee & Schickedanz 2007).

The PSRIP Phase 1 external evaluation showed gains in all components of reading with the exception of comprehension. As a result, after some research into teaching practices designed to scaffold and develop learners' comprehension skills through reading aloud, the design team decided to incorporate 'think alouds' into the PSRIP Shared Reading methodologies. This strategy is designed to expose learners to the internal thought processes of fluent readers, and to a variety of comprehension strategies, such as inferencing and making evaluations or judgments about text. During a think aloud, the idea is that teachers verbalise their thoughts while reading a text aloud, thereby exposing the invisible processes of comprehension. The purpose of this type of think-aloud strategy is to model how skilled readers monitor their own comprehension and ultimately, construct meaning from text.

In Phase 1 of the PSRIP, read-aloud texts were written to accompany images from existing DBE 'Big Book' story illustrations, images from the DBE Workbook, and posters. The idea was that teachers would read the text whilst learners looked at the selected illustrations. However, due to the format constraints, these texts were limited, and this was not an ideal solution. Later, when budget allowed for the inclusion of original stories, and after further research about reading comprehension, the decision was made to consciously use Shared Reading as a time to build learners' general (background) knowledge about a wide variety of topics by including a wider range of characters, settings, and plots than the original iteration allowed. It is well accepted that general knowledge of a topic makes comprehending text about the same concept easier (see, for example: Recht & Leslie 1988) – and inversely, that lack of background knowledge can make comprehension of text much more difficult.

The PSRIP Big Book stories were developed with the aim of increasing learners' background knowledge about a wide variety of topics, including a wide variety of socio-emotional concepts and issues, as well as countries, characters, and ideas from the wider world. Table 1 below provides an overview of the themes and stories developed for Grade 1 EFAL, together with the intended knowledge-development and targeted-comprehension strategy for each story.

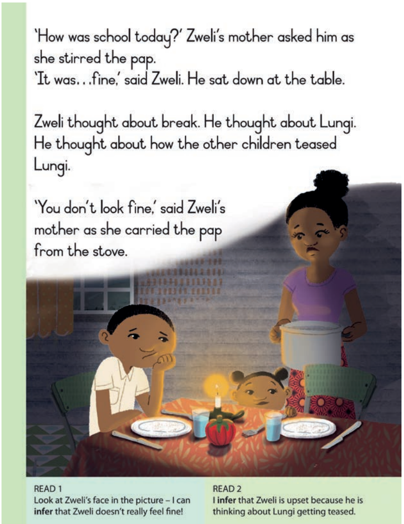
Table 1 Grade 1 PSRIP EFAL themes and shared reading stories (sample extract)

Grade and Term	Theme	Shared Reading Stories	General Knowledge Development	Comprehension Strategy
Grade 1, Term 1	We go to school	• Ben goes to school • Olwethu's first day	Both stories engage with the activities and social-emotional experience of attending school	• Search the text • Make connections
Grade 1, Term 1	My family	• Bongi waits • Tseko's new baby	• Exploring family dynamics and learning patience • Breaking down gender stereotypes and learning the skills required of an older sibling	• Visualise • Wonder / make inferences
Grade 1, Term 1	We play outside	• A very hot day • Spring Day splashes	• Experiencing different environments • A township tradition of splashing people on the first day of Spring	• Search the text • Visualise
Grade 1, Term 1	We have feelings	• Dan has a bad week • My name is Buhlebendalo	• The importance of planning, demonstration of growth mindset and resilience • The importance of name and identity, and learning to speak up for oneself	• Make connections • Make inferences

In the short term, a shallow level of implementation would mean teachers have prepared both the story and the think aloud, and are able to implement the strategy as per the core methodology. In the long term, deeper implementation would occur as teachers themselves become more familiar with the meta-language of comprehension, and the strategies needed to build these higher-level skills.

The most updated methodology for Shared Reading in the PSRIP then includes a 'pre-read', in which the teacher guides learners through a picture walk. In the context of EFAL, this helps expose learners to some of the vocabulary they will need to comprehend the text. In addition, the pre-read is focused on building the skill of prediction. In the first and second reads, the teacher reads the text out loud to learners, modelling fluency (accuracy and prosody) as she reads. In addition, the teacher uses the structured think-alouds detailed in the Read 1 and Read 2 blocks at the bottom of each page (Figure 1) to guide learners' thinking throughout the read-aloud. This helps learners to focus on key information and ultimately, be able to engage in discussion about the targeted questions at the end of the text.

Figure 1 **An example of a page showing the think-aloud text**



Further, during the first and second read, learners are exposed to the technical language of comprehension (i.e. 'Let's visualise' or 'I can infer') and are provided with concrete examples of what the strategies look like in practice. Finally, during the post-read activity learners demonstrate their understanding of the text through oral and written (in Grade 3 only) comprehension activities, including summarising the story, creating their own illustrations of the text, and dramatising the text, which includes recall.

While the Shared Reading methodologies may seem simple in theory, they embed a variety of best practices that are new to many South African classrooms as evidenced above. In addition, while the act of reading aloud may seem like a simple teaching task, it requires dedicated teacher preparation to work well. The teacher must read the story, understand the main purpose and lesson of the story, be able to read fluently and with expression to keep learners engaged, and know where to build interest, suspense, etc. Then, in addition, the teacher must include the act of thinking aloud. While the boxes at the bottom of the stories provide the content for teachers to use, to do this effectively, teachers need to understand the difference between reading aloud and thinking aloud, be able to differentiate their voices and actions so learners know when the teacher is reading versus thinking, keep a good pace, not distract learners from the plot of the story with too many words, and to use the key comprehension strategy words to build learners' technical vocabulary, etc. While at first the use of the Shared Reading methodologies may be mechanical, as time goes by and understanding builds, teachers are ideally able to implement this practice better and with increased conceptual understanding.

As a further example of the evolving design process, the PSRIP approach to teaching phonics and decoding has undergone a number of changes as a result of both participant feedback and ongoing research of evidence-based best practices for the teaching of reading (for example: Armbruster 2010; Ontario Human Rights

Commission 2022). At the beginning of the programme, a wider variety of both term- and grade-specific activities were included in the SLP methodologies, but through participant feedback, it was determined that a streamlined approach would lead to more consistent and accurate teaching of the phonics programme. This feedback led to the simplification of the phonemic awareness and phonic activities, in that the same set of activities – albeit with different sounds – are now used for an entire year.

Then, within the parameters of South Africa's National Curriculum Assessment Policy Statement (CAPS), the design team was conscious of the importance of building in both a systematic and explicit system for the teaching of English phonics. The programme design requires teachers to introduce a new phoneme to learners each week. Following this, throughout the rest of the week, a series of activities that align to evidence-based practices are completed; learners are required to identify, manipulate, blend, and segment new and previously learned phonemes and graphemes. And, as per the evidence, word building begins as soon as possible (National Reading Panel 2000).

While Group Guided Reading is designated in CAPS as an activity "to give learners individual attention in order to develop their comprehension and word attack skills" as per its whole-language origins as designed by Marie Clay and further developed by Fountas & Pinnell (Fountas et al. 2001), in the PSRIP SLPs, Group Guided Reading has been designated as a time to build learners' technical decoding skills. It is critical to note that the PSRIP design team has, over the course of the project, consciously shifted away from the original, intended origins of Group Guided Reading as a formal methodology, instead using the time allocation in CAPS for applying phonemic awareness and phonic knowledge to the technical reading of decodable texts. This focus has become more intentional and explicit over time. This means, for example, in the most recent PSRIP methodologies and training support, teachers are taught that the only word-attack strategy taught is that learners must sound unfamiliar words out, and that in the early stages, comprehension of text is not the main focus of the activity.

Accordingly, while early iterations of the PSRIP provided a core methodology to guide teachers through a standard sequence of activities to be completed with each small group, ongoing research and development throughout the implementation of the programme led the design team to acknowledge that the initial design was problematic and not aligned to decoding best practices (Ehri 1998; Hoffman et al. 2002; The Reading League 2022). Thus, teachers were directed to use a standard methodology with whichever reading materials they had access to – stipulating graded or levelled readers as the preferred text type (CAPS), or, if nothing else was available, texts from the DBE Workbook.

This early core methodology also promoted the usage of the five finger strategy mentioned in the Foundation Phase Home Language CAPS Document (2011). However, with the opportunity to revise and redesign the SLPs as part of the PSRIP Phase 3, the design team focused their attention on creating decodable texts (worksheets), which are aligned to the PSRIP phonics programme, allowing learners to apply their phonemic awareness and phonic knowledge to word and sentence reading. These worksheets include daily activities to practise reading decodable words, sight words (irregular pronunciation of graphemes, but still taught through sounding out), short decodable texts, and written comprehension activities. While originally these reading worksheets were designed to be used independently by the majority of the class while the teacher worked with a small group, the design team has since realised the importance of using these phonic-aligned, decodable texts during small group instruction as well.

4 Training and support

Ongoing training and support has been a critical component of the PSRIP from the outset; while the design and development of the classroom materials and resources described above is significant, there is ample evidence in the literature that materials and resources alone are insufficient to impact classroom change (Glewwe et al. 2001; Sabarwal et al. 2014). In the context of structured pedagogical interventions, expert-led training and professional classroom support (coaching) have been shown to impact the extent to which teachers are able to use the materials and resources that are provided effectively in their classrooms, and follow new teaching methodologies (Chakera et al. 2020; Alsofrom 2018). While a full and detailed description of the training programme is beyond the scope of this paper, this section aims to provide a high-level description of the design of the training programme.

As a sector programme, Foundation Phase subject advisors were expected to function as both trainers and coaches. The PSRIP's goal has been to capacitate subject advisors, to cascade the PSRIP-specific training to teachers, and to provide ongoing support (and not just monitoring) to their PSRIP teachers. Through experiences in previous interventions (GPLMS, EGRS I), the design team has learned that in the South African context, both trainers and coaches benefit from ongoing professional development to effectively support teachers; indeed, the need for the professional development of coaches is a theme within the wider literature (Atteberry & Bryk 2011; Bean 2014). As such, since 2017, PSRIP training has happened twice per year. Training for subject advisors is centralised, residential, and has a duration of three to four days per session. This training has been led by PSRIP master trainers – who are also trained prior to each of the PSRIP training sessions.

A few key principles have informed the design of all of the PSRIP trainings – including training of master trainers and subject advisors. Firstly, in line with both evidence and the design team's anecdotal experience, training was designed with the understanding that the professional development of teachers is unlikely to work when it is disconnected from everyday teaching practice, does not specifically align to curriculum implementation, or is too theoretical or decontextualised (Hayes 2000; Darling-Hammond et al. 2017). Based on this, initial training focused on changing practice through the methodologies in the structured learning programme. Only once subject advisors (and teachers) had time to practise the methodologies and see the benefits of using the SLPs, did training begin to include deeper theoretical and pedagogical knowledge.

Secondly, from the outset, the training has been designed not only to build technical skills, but to build emotional investment through relationship building. As Alsofrom (2018) argues, emotional engagement and trust-building are likely to be necessary (though insufficient) for creating the space to change teaching practices. In order to build emotional comfort and trust, for example, each of the PSRIP training sessions begins with a review and reflection session, where participants can genuinely share their experiences and reflect on successes and challenges, helping to build trusting relationships and to ensure that participants feel supported and heard. In addition, subject advisors are encouraged through the training process to be supportive resources for their teachers – and to help teachers feel comfortable asking questions about implementation challenges – rather than to simply monitor teachers' progress.

5 Learning from PSRIP experience

5.1 Focus on 80% of the system

The programme has an explicit concern with improving the teaching of reading in 80% of the school system where it remains a challenge: quintiles one to three, less wealthy, rural schools, which for historical and other structural reasons (low levels of district level support, inadequate school-based evaluation) do not have the resources or practices in place which enable reading success. Those schools in 20% of the system (quintiles four to five, urban or wealthier schools) do not necessarily need the resources offered through the PSRIP.

5.2 Use existing resources

Making use of existing district-level support services in the education system can be both a major constraint and a potential moment for innovation. South Africa has a total of approximately 554 advisors who are responsible for support to teachers in the Foundation and Intermediate Phases, across 17,442 public ordinary schools (DBE, 2021a). In the early grades advisors typically support the entire phase, whereas in the Intermediate Phase they support the teaching of English (FAL) as a subject. Their responsibilities include the implementation of multiple programmes at the same time and supporting schools in widely disparate community settings. There is considerable debate about the decision to rely on the existing corps of subject advisors and Foundation Phase specialists who are overstretched and under-resourced in most parts of the system.

It is not uncommon for schools in poorly resourced provinces and districts to report that they have not seen a subject advisor for the better part of a year. In that case, why rely on them when school-level professional support to teachers is a known factor in improving reading outcomes? PSRIP had to make reading improvement workable without the benefits of intensive school-level coaching – not by choice, but simply because it was not feasible in the existing institutional or financial parameters of the system. For example, while the programme materials are designed in line with term-by-term implementation of CAPS, four training sessions a year (face-to-face) are simply not practical, affordable, or an efficient use of teachers' time. As a result, training sessions are limited to two per year and further support to schools remains the responsibility of subject advisors and department heads (who also receive PSRIP training) – this is the norm in most new programmes rolled out across the sector. Critics have argued that this overall 'light-touch' approach will not achieve the necessary change in reading outcomes (Zenex, 2022). Herein lies the nub of the implementation challenge; what works in the context of well-resourced research studies must be replicable in an existing system with multiple binding constraints – of these, high personnel costs are probably the most difficult to overcome. Intensive school support cannot be achieved without changing the supply of professional support personnel on a massive scale.

5.3 Change management

PSRIP follows in the footsteps of numerous large-scale efforts to improve reading outcomes, including: Foundations for Learning (2008), Read to Lead (2015), Drop All and Read (2015). Changing the instructional core around teaching reading involves a complex change journey for teachers and those tasked with supporting their professional practice. Over time, subject advisors and teachers have acknowledged that the teaching of technical reading skills has improved significantly. As one subject advisor from the Western Cape Province noted in August 2019, at a nationwide review of the PSRIP, "I think the programme is all about the methodologies, and it's really a recipe that works. If the teachers implement this recipe and follow it step by step, they have huge success in their classes, and I'm very passionate about it". Another subject advisor from the North-West Province explained, "I tracked the learners that started this programme that are in Grade 6 now, and really when you look at the language, the level of language, the level of thinking, the level of writing, it's superb". While this evidence is anecdotal, it is representative of a major shift over time. A provincial coordinator from the Northern Cape Province referred to this in her reflection at the PSRIP review (August 2019): "I loved the way the trainers went off to find out any asked information and came back with feedback. I also felt like we were welcomed to talk and participate in the training – we felt heard". Being heard, being valued, being the driver of positive change is critical to the success and longevity of the PSRIP, in a context where change is often slow and difficult.

5.4 Systemic change

PSRIP is one piece of a reading ecosystem which is in midst of a major change cycle. There are undoubtedly gaps in the ecosystem which still need to be addressed. Applying Rogers' (1962) seminal work on how people and institutions react to change, it can be said that one third of schools are early adopters of an innovation like PSRIP, another third need help, and the last third will most likely never change. That challenge remains for managers of the system. At the very least, PSRIP has made a well-designed package of resources and methodologies available to the system – at sufficient scale – to provide a basis for that system-level transformation. It has also developed the capacity of subject advisors to support teachers and has enhanced pedagogical practices in relation to the teaching of reading.

6 Part II

6.1 Implementation evidence

In terms of teacher professional development, PSRIP was designed around the completion of a requisite period of training (twice yearly); completion of pre- and post-assessments at every session; and submission of a portfolio of evidence. These criteria are now standard practice in the sector and fulfil both a compliance requirement for statutory training funders (e.g. Sector Education and Training Authorities) and regulatory authorities (e.g. South African Council of Educators), which govern the award of continuous professional teacher development credits (CPTD).

The evidence from successive waves of pre- and post-assessments is one of the few channels providing programme developers with data on the short-term impacts of training and also enabling real-time adjustments in training content, quality, and delivery. Assessments were typically designed around two constructs: a) pre-existing knowledge of CAPS policy requirements (weighted at 40%), and b) training content (skills and knowledge – weighted at 60%). Assessments were conducted in face-to-face training sessions, marked, and moderated independently. Subject advisors and teachers were assessed using customised instruments, with advisors being expected to perform at a higher level.

Table 2 indicates the coverage of the PSRIP over the period 2017-2020 which by South African standards provides a significant sample size from which to draw. By 2020, the programme had reached just under 30,000 primary school teachers in terms of the numbers trained, provided with teaching materials, and supported.

Table 2 Number of subject advisors and teachers reached 2017-2020 by PSRIP

Period: 2017–2020: PSRIP 1, 2, 3 Cumulative Reach					
Province	**Schools**	**FP SAs**	**IP SAs**	**FP Teachers**	**IP Teachers**
EC	1,509	90	44	2,816	1,715
WC	589	56	31	2,015	1,101
NC	451	39	24	1,234	621
GP	678	66	35	1,623	445
MP	551	45	12	2,032	476
NW	704	43	28	1,757	1,003
LP	1,024	56	43	2,801	1,653
KZN	1,140	69	43	3,533	1,528
FS	596	58	22	2,192	846
Totals	**7,240**	**522**	**282**	**20,003**	**9,388**

The following data are extracted from reporting on the results of pre- and post-assessments in PSRIP II and III of the programme conducted by the Monitoring and Quality Assurance (MQA) team of the NECT and external analysts.

Table 3 and Figure 2 illustrate the degree of change in test outcomes for Foundation Phase subject advisors. It is difficult to ignore the low base from which these scores started before improving, and the implications this may have for this critical group of professionals. None of the provinces achieved above 50% at pre-test stage. There are also anomalies in provinces, showing a decline at post-assessment stage (Table 9). The PSRIP relies fundamentally on subject advisors being sufficiently skilled and knowledgeable to train and support management teams and teachers at school level.

This is in stark contrast to the results attained by Intermediate Phase subject advisors who started from a much stronger base in Phase II. This cohort also made significantly better gains in the short term (Figures 3, 4, and 5). It is notable that the under-performance of Foundation Phase subject advisors is mirrored in the performance of 653 Foundation Phase teachers (Table 4, Figure 6). In Table 5 and Figure 7, Intermediate Phase teachers in a similarly large sample (813) show similar patterns of underperformance at pre-test stage.

In Phase III, Table 6 and Figure 8 show a marked uptick in the performance of Foundation Phase advisors, while Intermediate Phase advisors are even stronger (Table 7). Tables 8 and 9 also show positive changes, this time in teachers' performance – in both Foundation and Intermediate phases – across large samples of 572 and 723 teachers respectively.

Table 3 Summary of Foundation Phase subject advisor test scores (Phase II) (*n*=119)

Province	Pre/20 Sec A	Pre% Sec A	Post/20 Sec A	Post% Sec A	Pre/30 Sec B	Pre% Sec B	Post/30 Sec B	Post% Sec B	Pre Tot %	Post Tot %
EC	7.1	35.6%	8.0	40.2%	12.7	42.2%	19.7	65.7%	**39.6%**	**55.5%**
FC	7.1	35.4%	7.4	37%	9.5	32.1	16.6	55.5%	**33.4%**	**48.1%**
GP	7.6	37.9%	8.9	44.7%	11.1	37%	16.3	54.3%	**37.4%**	**50.4%**
KZN	7.2	35.9%	7.1	35.7%	12.0	40%	16.7	55.5%	**38.3%**	**47.6%**
LP	7.3	36.7%	8.1	40.4%	11.2	37.2%	14.4	47.9%	**37.0%**	**44.9%**
MP	6.7	33.4%	8.0	40.2%	10.0	33.5%	14.4	47.9%	**33.4%**	**44.8%**
NC	7.5	37.7%	7.3	36.3%	12.4	41.2%	16.3	54.5%	**39.8%**	**47.2%**
NW	8.7	43.3%	8.1	40.8%	14.9	49.5%	16.7	55.6%	**47%**	**49.7%**
WC	9.1	45.4%	8.2	40.9%	12.1	40.3%	18.3	61.1%	**42.3%**	**53%**
Nat	**7.6**	**37.9%**	**7.9**	**39.5%**	**11.8**	**39.2%**	**16.6**	**55.3%**	**38.7%**	**49%**

Figure 2 Progression per province pre- and post-test

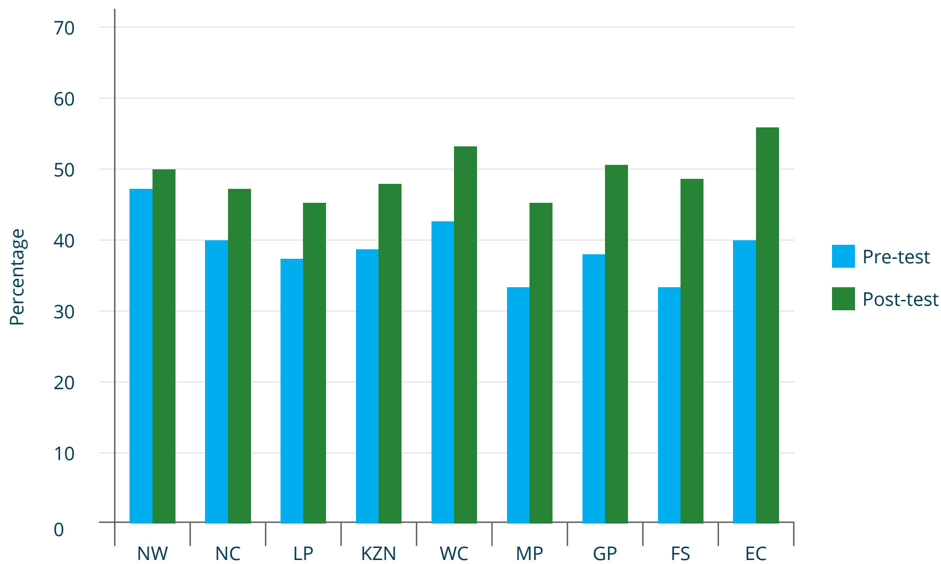


Figure 3 Intermediate Phase subject advisors (IP) pre- and post-test Scores (n= 82)

Province	Pre /20 Sec A	Pre % Sec A	Post /20 Sec A	Post % Sec A	Pre /20 Sec B	Pre % Sec B	Post /20 Sec B	Post % Sec B	Pre /10 Sec C	Pre % Sec C	Post /10 Sec C	Post % Sec C	Pre Tot %	Post Tot %
EC	9.5	47.5%	13	65.0%	8	40.0%	11	55.0%	5	50%	8	80%	**45%**	**64%**
FS	12.7	63.4%	12.4	62.0%	5.4	27%	10.6	53.0%	2.8	28%	7.2	72%	**41.6%**	**60.4%**
GP	12.4	62.0%	13.9	69.5%	6.3	31.5%	13.6	68.0%	3.2	32%	6.6	66%	**43.6%**	**68.2%**
KZN	12.6	63.0%	12.2	61.0%	5.4	27.0%	11.4	57.0%	1.7	17%	5.8	58%	**39.4%**	**58.8%**
LP	11.3	56.5%	11	55.0%	4.2	21.0%	4.9	24.5%	1	10%	2.2	22%	**33.0%**	**36.2%**
NC	10.6	53.0%	12.7	63.5%	5.3	26.5%	14.4	72.0%	3.1	31%	6.1	61%	**38.0%**	**66.4%**
NW	11.9	59.5%	11.2	56.0%	5.7	28.5%	7.7	38.5%	2.3	23%	4.1	41%	**39.8%**	**46.0%**
WC	11.7	58.5%	12.1	60.5%	4.6	23.0%	6.2	31.0%	1.8	18%	1.9	19%	**36.2%**	**40.4%**
Nat		**57.9%**		**61.6%**		**28.06%**		**49.8%**		**26.0%**		**52.4%**	**39.2%**	**54.1%**

Figure 4 Provincial mean scores: Intermediate Phase subject advisors (*n*=82)

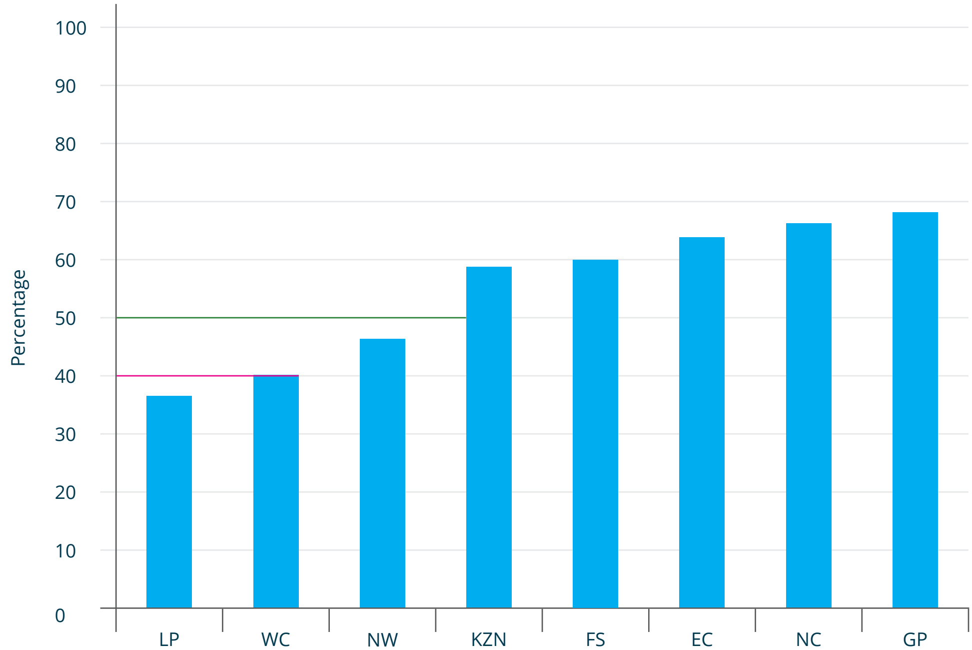


Figure 5 Subject Advisors Intermediate Phase: progression per province (*n* = 653, randomly selected)

Table 4 Foundation Phase teachers: pre- and post-test scores

Province	Pre/14 Sec A	Pre % Sec A	Post/14 Sec A	Post % Sec A	Pre/21 Sec B	Pre % Sec B	Post/21 Sec B	Post % Sec B	Pre Tot %	Post Tot %
EC	4.5	31.9%	5.4	38.8%	5.8	27.5%	9.8	46.9%	**29.3%**	**43.6%**
FC	4.6	32.6%	5.9	42.5%	6.2	29.7%	9.5	45.2%	**30.8%**	**44.1%**
GP	4.3	30.8%	5.4	38.3%	8.6	40.7%	11.6	55.3%	**36.8%**	**48.5%**
KZN	3.6	25.4%	4.6	32.9%	4.9	23.4%	8.9	42.4%	**24.2%**	**38.6%**
LP	5.3	37.9%	6.9	49.6%	9.2	43.7%	12.6	60.0%	**41.4%**	**55.8%**
MP	4.5	32.1%	5.2	37.4%	7.0	33.2%	10.3	49.1%	**32.8%**	**44.5%**
NW	5.6	40.0%	5.9	42.1%	8.0	38.3%	10.9	51.9%	**39.0%**	**48.0%**
NC	4.7	33.2%	5.9	42.4%	6.6	31.3%	12.2	58.1%	**32.1%**	**51.8%**
WC	4.0	28.3%	5.5	39.0%	6.1	29.2%	10.1	48.1%	**28.8%**	**44.5%**
Nat	**4.5**	**32.4%**	**5.6**	**40.1%**	**6.7**	**32.1%**	**10.5**	**50.1%**	**32.2%**	**46.1%**

Figure 6 Foundation Phase teachers: pre- and post-test scores (*n* = 813)

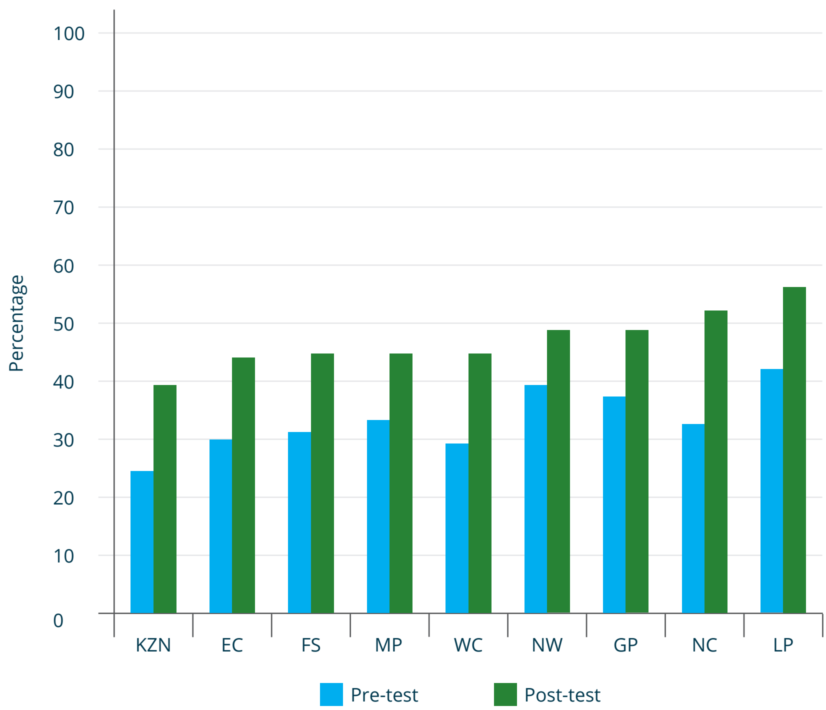


Table 5 Intermediate Phase teachers: pre- and post-test scores

Province	N	Pre% Sec A	Post% Sec A	Pre% Sec B	Post% Sec B	Pre% Total Score	Post% Total Score
EC	177	55.0%	59.5%	32.5%	53.9%	44.6%	56.9%
FC	62	55.4%	52.9%	27.0%	39.9%	42.35%	46.9%
GP	85	57.3%	57.9%	35.5%	46.35%	47.2%	52.5%
KZN	115	57.3%	58.5%	34.3%	61.6%	46.7%	59.9%
LP	92	55.0%	53.5%	30.2%	39.7%	43.6%	47.1%
MP	40	51.9%	53.1%	24.7%	43.3%	39.3%	48.6%
NW	69	52.7%	54.5%	44.9%	45.2%	49.15%	50.2%
NC	77	56.8%	55.8%	32.7%	55.8%	45.75%	55.8%
WC	96	52.6%	54.3%	26.8%	54.0%	40.75%	54.2%
Nat	813	55.2%	56.3%	32.4%	50.5%	44.75%	53.6%

Figure 7 Intermediate Phase teachers: pre- and post-test scores (*n*=117)

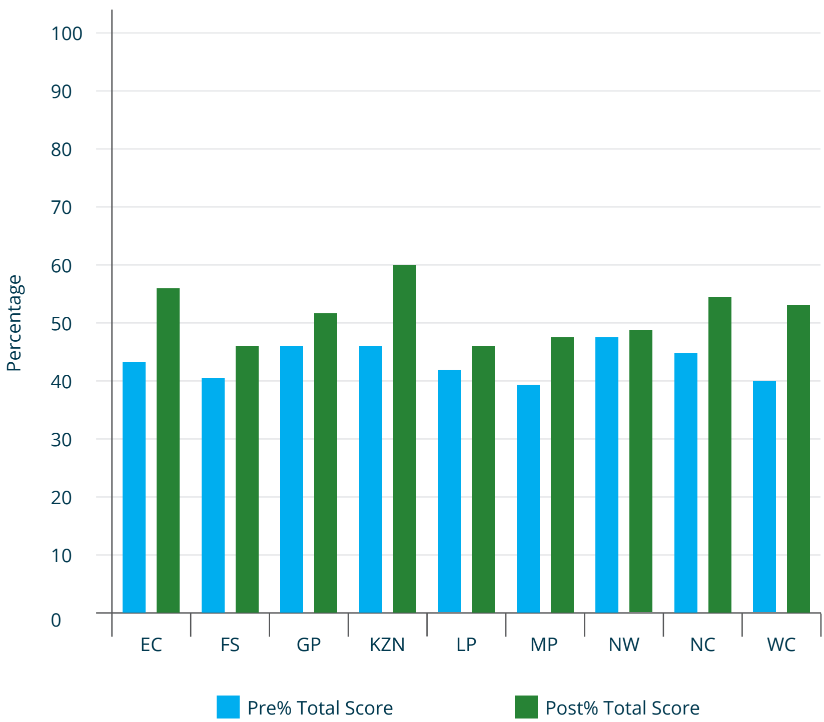


Table 6 Foundation Phase subject advisors: pre- and post-test scores (Phase III)

Province	Pre % Sec A	Post % Sec A	Change (percentage point p.p.)	Pre % Sec B	Post % Sec B	Change (percentage point p.p.)	Pre % Total Score	Post % Total Score	Change (percentage point p.p.)
EC (*n* = 19)	71	80	+9p.p.	69	75	+6p.p.	**70**	**77**	**+7p.p.**
FS (*n* = 17)	66	79	+13p.p.	60	71	+9p.p.	**63**	**74**	**+9p.p.**
KZN (*n* = 17)	79	88	+9p.p.	68	82	+14p.p.	**73**	**84**	**+11p.p.**
LP (*n* = 17)	84	91	+7p.p.	66	79	+13p.p.	**74**	**84**	**+10p.p.**
MP (*n* = 12)	74	80	+6p.p.	69	79	+10p.p.	**71**	**79**	**+8p.p.**
NC (*n* = 3)	83	89	+6p.p.	78	87	+9p.p.	**80**	**88**	**+8p.p.**
NW (*n* = 15)	76	84	+8p.p.	70	75	+5p.p.	**73**	**79**	**+6p.p.**
WC (*n* = 17)	69	82	+13p.p.	52	75	+23p.p.	**59**	**78**	**+19p.p.**
Nat	74.5	83.3	+8.8p.p.	65.1	76.7	+11.2p.p.	68.9	79.2	+10.3p.p.

Figure 8 Foundation Phase subject advisors: pre- and post-test scores (Sec A & B) (*n*=148)

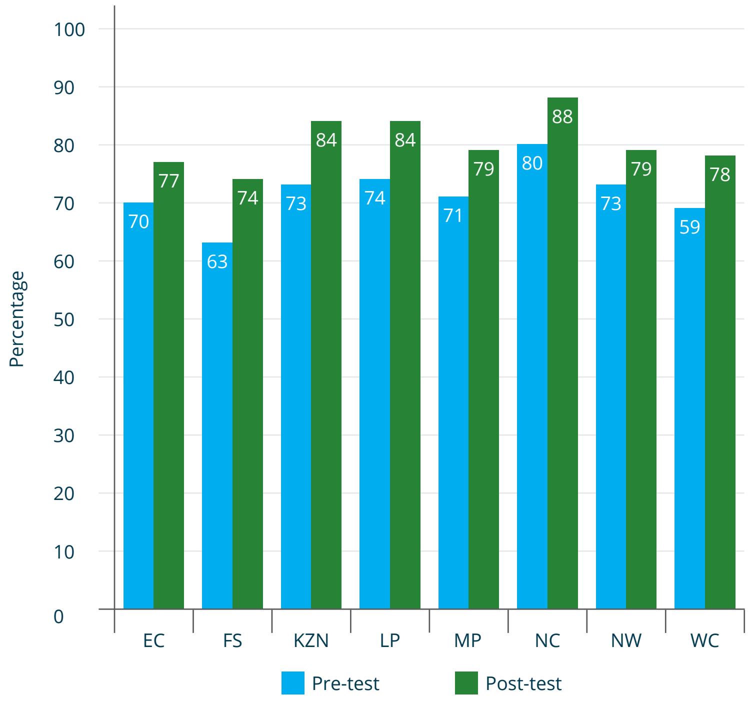


Table 7 Intermediate Phase subject advisors pre- and post-test scores (Phase III)

Province	Pre% Sec A	Post% Sec A	Change (percentage point p.p.)	Pre% Sec B	Post% Sec B	Change (percentage point p.p.)	Pre-Total %	Post-Total %	Change (percentage point p.p.)
EC (*n* = 24)	76.2	79.9	+3.7p.p.	22.8	37.7	+14.9p.p.	**38.0**	**49.7**	**+11.7p.p.**
FS (*n* = 10)	76.4	84.6	+8.2p.p.	31.3	44.2	+12.9p.p.	**46.7**	**58.9**	**+12.2p.p.**
GP (*n* = 19)	82.0	85.3	+3.3p.p.	26.5	53.5	+27.0p.p.	**43.9**	**67.7**	**+23.8p.p.**
KZN (*n* = 21)	77.6	78.4	+0.8p.p.	20.9	38.0	+17.1p.p.	**38.0**	**49.5**	**+11.5p.p.**
LP (*n* = 22)	74.5	76.3	+1.8p.p.	25.5	45.5	+20.0p.p.	**40.1**	**53.9**	**+13.6p.p.**
MP (*n* = 6)	65.5	82.1	+16.6p.p.	20.6	35.2	+14.6p.p.	**33.3**	**49.2**	**+15.9p.p.**
NC (*n* = 12)	69.0	74.4	+5.4p.p.	16.7	38.1	+21.4p.p.	**31.7**	**49.1**	**+17.4p.p.**
NW (*n* = 18)	77.6	77.6	0p.p.	20.0	35.3	+15.3p.p.	**36.4**	**47.2**	**+10.8p.p.**
WC (*n* = 16)	82.6	83.0	+0.4p.p.	28.4	46.1	+15.7p.p.	**43.6**	**60.4**	**+16.8p.p.**
Nat	**76.7**	**79.9**	**3.2p.p.**	**23.6**	**41.3**	**+17.7p.p.**	**39.4**	**53.4**	**+14.0p.p.**

Table 8 Foundation Phase teacher pre- and post-test scores (Phase III) (*n* = 572)

Province	Pre% Sec A	Post% Sec A	Change (percentage point p.p.)	Pre% Sec B	Post% Sec B	Change (percentage point p.p.)	Pre-Total %	Post-Total %	Change (percentage point p.p.)
EC (*n* = 78)	45.9%	46.2%	0.3p.p.	50.4%	57.2%	6.8p.p.	**49%**	**53%**	**4p.p.**
FS (*n* = 88)	37.0%	44.2%	7.2p.p.	38.7%	49.1%	10.4p.p.	**38%**	**47%**	**9p.p.**
KZN (*n* = 76)	39.3%	45.2%	5.9p.p.	37.9%	51.1%	13.2p.p.	**38%**	**49%**	**11p.p.**
LP (*n* = 74)	43.5%	45.4%	1.9p.p.	43.0%	54.1%	11.1p.p.	**43%**	**51%**	**8p.p.**
MP (*n* = 45)	32.8%	51.1%	18.3p.p.	48.1%	66.9%	18.8p.p.	**42%**	**61%**	**19p.p.**
NC (*n* = 49)	42.5%	62.1%	19.6p.p.	50.1%	69.7%	19.6p.p.	**47%**	**67%**	**20p.p.**
NW (*n* = 79)	45.8%	55.3%	9.5p.p.	44.6%	51.5%	6.9p.p.	**45%**	**53%**	**8p.p.**
WC (*n* = 83)	39.5%	54.3%	14.8p.p.	31.2%	53.3%	22.1p.p.	**34%**	**54%**	**20p.p.**
Nat	**41.1%**	**49.8%**	**8.7p.p.**	**42.2%**	**55.2%**	**13.0p.p.**	**41.6%**	**53.3%**	**11.7p.p.**

Table 9 Intermediate Phase teacher pre- and post-test scores (*n* = 723)

Province	Pre% Sec A	Post% Sec A	Change (percentage point p.p.)	Pre% Sec B	Post% Sec B	Change (percentage point p.p.)	Pre-Total %	Post-Total %	Change (percentage point p.p.)
EC (*n* = 38)	62.8	69.4	6.6p.p.	12.0	30.9	18.8p.p.	**32.3**	**46.3**	**14p.p.**
FS (*n* = 68)	67.5	71.3	3.8p.p.	13.5	27.8	14.3p.p.	**35.1**	**45.3**	**10.2p.p.**
KZN (*n* = 130)	70.5	71.1	0.6p.p.	18.1	35.0	16.9p.p.	**39.0**	**49.4**	**10.4p.p.**
LP (*n* = 125)	67.7	69.7	2.0p.p.	11.1	18.2	7.1p.p.	**33.8**	**38.8**	**5.0p.p.**
MP (*n* = 24)	65.0	71.6	6.6p.p.	17.3	27.1	9.8p.p.	**36.4**	**44.9**	**8.5p.p.**
NC (*n* = 54)	74.4	74.1	−0.3p.p.	23.9	36.6	12.5p.p.	**44.1**	**51.6**	**7.5p.p.**
NW (*n* = 88)	67.9	70.8	2.9p.p.	16.0	28.9	12.9p.p.	**36.8**	**45.6**	**8.8p.p.**
WC (*n* = 96)	73.2	74.3	1.1p.p.	19.0	35.7	16.7p.p.	**40.7**	**51.1**	**10.4p.p.**
Nat	**68.4**	**71.2**	**2.8p.p.**	**15.6**	**29.8**	**14.2p.p.**	**36.7**	**46.4**	**9.7p.p.**

7 Findings and analysis

External evaluations were also conducted on each phase of the PSRIP with the routine focus on programme outcomes (NECT, 2018). The data used here are indicative of what is typically available to provincial and national managers on the short-term outcomes of professional development programmes for subject advisors and teachers. Provincial education departments are by far the largest providers of in-service professional development, and therefore control the bulk of available resources in the system. Whilst PSRIP may be relatively large in scale, it is dwarfed by provincial teacher development budgets. Furthermore, it is a small though significant response to the challenge of systematically reaching 90,000 (primary and secondary) teachers in a province like KwaZulu-Natal, the largest employer in the system.

What do the data tell us at face value? How can the data be used? Some observations:

7.1 Management of professional development

- Pre- and post-assessments are necessary and useful if designed and used well. They require time and expertise to develop and willingness on the part of participants to improve their capabilities.
- Whilst they may have been driven initially by compliance requirements, there are signs that 'value for money' considerations are increasing the level of attention to programme outcomes. A recently completed report by the Department of

Performance Monitoring and Evaluation (DPME 2021) is a useful bellwether.

- Subject advisors and teachers were initially reluctant to participate in pre- and post-assessments until being assured that the results were important to improving the programme's delivery and management. The fear of punitive judgement is a major deterrent.
- Consequently, it is common knowledge that self-administered assessments as part of professional development programmes also have very poor uptake in the system.
- Assessments have only become the norm in the past six years and take time to become routinised – i.e. used consistently, non-judgementally, and with an emphasis on support. Ideally, the assessments should be part of a coherent data-driven system of professional development.
- Evaluation of professional development programmes – particularly those designed and delivered in-house by provincial departments – is not required or routinised in the sector. This creates a significant gap in the evidence base with respect to what the needs are with respect to design, planning, impact, and value for money.

7.2 Results

- The data provide useful insights into the profile of participants over two cycles of implementation and strongly indicate patterns of results.
- From the initial data, it is fair to say that many education professionals responsible for teaching early grade reading are starting from a low base of policy knowledge and methodological skills. The pattern of under-performance of Foundation Phase subject advisors and teachers attests to this issue. For the programme to reach its goals, subject advisors need to be performing consistently at a much higher level, in order to ensure that they maintain programme fidelity.
- However, it is clear that performance on pre- and post-assessments for both subject advisors and teachers is showing positive change over time and at significant scale.
- Subject advisors' responses also indicate growing confidence in the use of their new methodologies.
- Whilst there may be challenges to the construct validity of policy questions, the data show serious inadequacies in knowledge of CAPS both at subject advisor and teacher level, with the exception of the Intermediate Phase, which should not be a problem at this stage in policy implementation.
- There are significant trends in the under-performance of some provinces which need direct attention, especially where district-level support to teachers is low and large numbers of teachers still need to be reached.
- The outcomes on the latter part of the assessments (Section B) are the more important since they are based on the training content. Overall, the data show positive outcomes but these changes need to be translated into practice; reinforced by reflection and evaluation; and supported by heads of department. Can these changes be sustained in the absence of effective school-level support by advisors or coaches?
- More detailed item analysis showed that the teaching of English phonics was a recurring problem area. Exploring the issue in more detail is beyond this paper and

this is an issue, which has been extensively addressed in programme-design updates.
- On a related note, PSRIP training also covers the use of Early Grade Reading Assessment (EGRA) in English. However, there is no systematic monitoring of how and when EGRA data is collected or used by teachers, nor is there a routinised evaluation of the data at school or district level.
- The programme and the sector need more detailed and longer-term outcome measures of professional development programmes. Better data from programmes across the spectrum would strengthen efforts at all levels of the system in planning, design and delivery.

8 Conclusion and way forward

There is a strong consensus that recent practice in initial teacher training has badly neglected the core skills in teaching reading particularly in the early grades. The PSRIP is one strategy to address that systemic challenge within the limits of what is doable at scale through continuous teacher professional development. A consensus around strengthening the instructional core in the early grades through evidence informed approaches has taken hold, albeit starting from a low base, and in a system which is recovering from significant learning losses due to the Covid-19 pandemic, notably in literacy and mathematics. However, good quality teaching resources and sufficiently tested methodologies are now available freely and widely to any teacher who needs them (NECT 2021). Advisors, teachers, and managers of the sector have demonstrated a positive appetite for changing teaching practices. The challenge lies in intentional and sustained support of these changes in teaching practice.

The programme has yielded important implementation evidence on teacher professional development in teaching reading, which could be applied to other initiatives in the sector. These data need to be used in conjunction with reliable learner outcome data from EGRA, which is currently lacking in the sector, at district and school levels. Teachers would be far better supported if they had reliable data on reading outcomes in their school, circuit, or district against which to benchmark their learners.

Progress towards evidence-informed normative standards in teacher development on reading, including methodologies, routines, and expected outcomes, is critical to better reading outcomes in basic education. Important efforts are in process to revitalise and reshape the reading ecosystem that should be supporting schools and communities, and this cannot be achieved without better quality, well planned, and evaluated teacher professional development.

Acknowledgement

This papers draws on data collected and analysed by the NECT's Monitoring and Quality Assurance (MQA) team including external consultants. Images are provided by ClassAct, a PSRIP implementing partner.

References

Alsofrom, K. (2018). How and why coaching works to improve teaching practices in the EGRS II? An examination of causal mechanisms. Report to the Early Grade Reading Study (EGRS).

Evans, M. & Saint-Aubin, J. (2005). What children are looking at during shared storybook reading: Evidence from eye movement monitoring. *Psychological science*. 16(11), 913–920.

Armbruster, B. B. (2010). *Put reading first: Kindergarten through Grade 3*. Washington, DC: Diane Publishing.

Atteberry, A., & Bryk, A. S. (2011). Analyzing teacher participation in literacy coaching activities. *The Elementary School Journal*. 112(2), 356–382.

Bean, R. M. (2014). *The power of coaching: Improving early grade reading instruction in developing countries*. Washington, DC: USAID.

Chakera, S., Haffner, D., & Harrop, E. (2020). Structured pedagogy: For real-time equitable improvements in learning outcomes. Working Paper 10701, UNICEF. Available at: https://www.unicef.org/esa/media/7511/file/ESA-Structured-Pedagogy-2020.pdf (Accessed 14 November 2022).

Conn, K. (2014). *Identifying Effective Education Interventions in Sub-Saharan Africa: A meta-analysis of rigorous impact evaluations.* Ann Arbor, MI: ProQuest.

Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2017). *Effective teacher professional development.* Palo Alto, CA: Learning Policy Institute.

Dehaene, S. (2009). *Reading in the brain*. New York: Penguin Books.

Department of Basic Education (2021). Rapid Review of Continuous Professional Development for Teachers in South Africa. Pretoria: DBE.

Department of Basic Education (2021a). School Masterlist Data. Available at: https://www.education.gov.za/Programmes/EMIS/EMISDownloads.aspx. (Accessed 15 November 2022)

DeMonte, J. (2013). High-quality professional development for teachers: Supporting teacher training to improve student learning. Center for American Progress.

DPME (Department for Planning, Monitoring and Evaluation) (2021). Rapid Review of Continuous Professional Development for Teachers in South Africa (unpublished)

Ehri, L. C. (1998). Grapheme–phoneme knowledge is essential for learning to read words in English. *Word recognition in beginning literacy*. 3 (40), 100–104.

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.

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. (2018). The education triple cocktail: System-wide instructional reform in South Africa. Cape Town: UCT Press/Juta and Company (Pty) Ltd.

Fleisch, B., & Dixon, K. (2019). Identifying mechanisms of change in the early grade reading study in South Africa. *South African Journal of Education*, 39(3).

Fountas, I. C., Pinnell, G. S., & Le Verrier, R. (2001). *Guided reading*. Portsmouth, NH: Heinemann.

Glewwe, P., Jacoby, H. G., & King, E. M. (2001). Early childhood nutrition and academic achievement: a longitudinal analysis. *Journal of Public Economics*, 81(3), 345–368.

Gold, J., & Gibson, A. (2001). Reading aloud to build comprehension. *Reading Rockets*, 32(7), 14–21.

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

Hayes, D. (2000). Cascade training and teachers' professional development. *ELT journal*, 54(2), 135–145.

Hoffman, J. V., Sailors, M., & Patterson, E. U. (2002). Decodable texts for beginning reading instruction: The year 2000 basals. *Journal of Literacy Research*, 34(3), 269–298.

Kilpatrick, D. A. (2015). *Essentials of assessing, preventing, and overcoming reading difficulties*. Hoboken, NJ: John Wiley & Sons.

Kim, Y.-S. G., & Davidson, M. (2019). Promoting successful literacy acquisition through structured pedagogy: Global Reading Network Critical Topics Series. Prepared by University Research Co., LLC. (URC) under the Reading within Reach (REACH) initiative for USAID's Building Evidence and Supporting Innovation to Improve Primary Grade Assistance for the Office of Education (E3/ED). Available at: www.edu-links.org (Accessed 14 November 2022).

Kim, Y. S. G. (2020). Toward integrative reading science: The direct and indirect effects model of reading. *Journal of Learning Disabilities*, 53(6), 469–491.

McGee, L. M., & Schickedanz, J. A. (2007). Repeated interactive read-alouds in preschool and kindergarten. *The Reading Teacher*, 60(8), 742–751.

Mugambi, F. N. (2015). Exploring reading habits and academic success in rural Kenya. *IFLA Journal*, 41(4), 353–363.

NECT (2019). Primary School Reading Improvement Programme. Available at: https://nect.org.za/publications/case-studies/psrip_booklet.pdf (Accessed 14 November 2022).

NECT (2021). Primary School Reading Improvement Programme – materials for teachers. Available at: https://nect.org.za/materials/psrip (Accessed 14 November 2022).

National Reading Panel (US), (2000). Report of the National Reading Panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups. National Institute of Child Health and Human Development, National Institutes of Health.

Ontario Human Rights Commission (2022). Right to Read: Public inquiry into human rights issues affecting students with reading disabilities. Ontario Human Rights Commission Approved by the OHRC: January 27, 2022 © February 2022, Government of Ontario.

Piper, B., Destefano, J., Kinyanjui, E. M., & Ong'ele, S. (2015). Scaling up successfully: Lessons from Kenya's Tusome national literacy program. *Journal of Educational Change*, 19(3), 293–321.

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

Pretorius, E., & Naude, H. (2002). A culture in transition: Poor reading and writing ability among children in South African townships. *Early Child Development and Care*, 172(5), 439–449.

Recht, D. R., & Leslie, L. (1988). Effect of prior knowledge on good and poor readers' memory of text. *Journal of Educational Psychology*, 80(1), 16.

Rogers, E.M. (1962). *Diffusion of Innovations*. New York, NY: The Free Press.

Sabarwal, S., Evans, D., & Marshak, A. (2014). The permanent input hypothesis: The case of textbooks and (no) student learning in Sierra Leone. World Bank Policy Research Working Paper, (7021).

Seidenberg, M. S., & Borkenhagen, M. C. (2020). The Reading League Journal. Available at: https://seidenbergreading.net/wp-content/uploads/2020/03/Reading-Science-and-Educational-Practice-Some-Tenets-for-Teachers.pdf (Accessed 14 November 2022).

Snilstveit, B., Stevenson, J., Menon, R., Phillips, D., Gallagher, E., Geleen, M., & Jimenez, E. (2016). *The impact of education programmes on learning and school participation in low-and middle-income countries*. London: International Initiative for Impact Evaluation.

Taylor, N. (2019). Inequalities in teacher knowledge in South Africa. In Spaull, N. & Jansen, J. (Eds), *South African Schooling: The Enigma of Inequality*. Springer: Cham.

Tella, A., & Akande, S. O. (2007). Children reading habits and availability of books in Botswana primary schools: Implications for achieving quality education. 2030 Reading Panel (2022), Background report.

Zenex Foundation (2022). NECT Early Grade Learning Programme Evaluation. Available at: https://www.zenexfoundation.org.za/wp-content/uploads/2022/03/NECT-Learning-Brief.pdf (Accessed 14 November 2022).

05

The science and magic of reading in action: Lessons learned from the Room to Read Literacy Programme in South Africa

KERRI THOMSEN, VANESSA FRANCIS & CATHERINE NGWANE

KEYWORDS
intervention, reading instruction, libraries, book publishing, habit of reading, reading skills, coaching, data-based decision making

Abstract

Room to Read has implemented its Literacy Programme in 880 schools in Limpopo, Mpumalanga, Gauteng, KwaZulu-Natal, and Eastern Cape since 2007. This chapter will use organisational documents, employee focus groups, and the authors' experiences to answer the question: having spent more than a decade addressing the acquisition of the skill and habit of reading in South Africa, what are the key lessons learned about implementing a literacy intervention? The lessons focus around key areas of materials and resources, teacher training and support, data-based decision-making, and family, community, and government engagement. The chapter concludes with a look forward to the challenges and opportunities for scale and systems change.

1 Introduction

Despite investments to improve education, only 22 per cent of Grade 4 children in South Africa could read for meaning in any language on the Progress in International Reading Literacy Study, that is, achieved the PIRLS Low International Benchmark, in 2006 and 2016 (Mullis et al. 2007; Mullis et al. 2017). Seeing a need in South Africa, Room to Read, an international non-profit organisation focused on literacy and gender equality in education, founded its South Africa office in 2006 with the goal of improving literacy in low-achieving districts.

The Room to Read Literacy Programme supports the development of lifelong, independent readers, which the organisation defines as readers who read often, voluntarily, and with enjoyment. This requires a focus on both literacy skills and the habit of reading (Brandt et al. 2021). The programme theory of change is that investing in material resources, training, and coaching for educators, monitoring and evaluation, and stakeholder engagement will produce lifelong learners with the skills and habit of reading, who are better prepared to make positive change. After 20 years of research- and data-driven programmatic changes, Room to Read seeks avenues for systemic change and opportunities for scale.

In South Africa, the organisation began to establish school libraries in 2007 and to conduct book publishing activities in 2008. In 2012, it added classroom literacy instruction alongside its school libraries programming. Following this, in 2016 Room to Read revised its instruction and library programming into a comprehensive Literacy Programme. As the organisation demonstrated the successes of the programme in schools, with gains over comparison schools in multiple subtests, including oral reading fluency (Room to Read 2014; Rigole & Phaweni 2018), it also continued to amend the programme for greater efficacy. In 2019, an in-school coaching programme with Heads of Department (HODs) was added. Since 2007, Room to Read South Africa has supported 692,790 learners in 880 schools (Table 1).

Table 1 Room to Read literacy projects in South Africa

	Provinces	Number of new schools	Number of new learners	Type of project school
2007-2011	Eastern Cape, Mpumalanga, Gauteng, Limpopo	292	205,715	Library only
2012-2013	Limpopo, Mpumalanga	116	82,581	Library only
		57		Instruction added to Library project schools
2014-2015	Mpumalanga, Gauteng, Limpopo	66	73,863	Comprehensive Literacy Programme
2016-2018	Limpopo, Mpumalanga	159	154,891	Improved Comprehensive Literacy Programme
2019-2021	Limpopo, Mpumalanga, KwaZulu-Natal	190	175,740*	Current Comprehensive Literacy Programme (including HOD Model)

Source: *2020 Global Indicator Report (Room to Read 2022) and 2021 estimates (Room to Read 2021)*

**At the time of writing, the number of new learners in 2021 was being finalised. This table includes an approximate number.*

1.1 Literacy Programme description

The Room to Read Literacy Programme currently supports primary schools, over a period of three years, i) to develop the home language[1] literacy skills of Grade 1 and 2 learners and ii) to encourage the habit of reading among all the primary school learners. Grade 1 and 2 teachers implement a structured pedagogy approach during their usual literacy time allocated in the school timetable. Schools are requested to ensure a compulsory 30-minute library period in the timetable, in which language teachers lead a reading activity and manage checking out books with the help of community volunteers. Table 2 details the inputs for the programme.

Table 2 Programme Inputs

<table>
<tr><th colspan="2">Dosage per Learner</th><th colspan="2">Ratio of Coaches to Schools</th></tr>
<tr><td colspan="2">Instruction time per day (Grades 1 to 3): 45 minutes
Library period per week (Grades R to 7): 30 minutes
Undesignated library time per week: 60 minutes</td><td colspan="2">One coach to ten schools</td></tr>
<tr><th colspan="4">Inputs</th></tr>
<tr><th colspan="2">Components & Activities</th><th colspan="2">Amount over 3 project years</th></tr>
<tr><td rowspan="9">Professional Development & Support</td><td rowspan="3">Literacy Instruction Training for teachers and Heads of Department (HODs)</td><td colspan="2">19 days (10, 6, 4) for Grade 1*
16 days (10, 6) for Grade 2
Six days (6) for Grade 3 (planned for 2024)</td></tr>
<tr><td colspan="2">Six days (3, 2, 1) for Foundation Phase (FP) HODs</td></tr>
<tr><td colspan="2">Three (3) days facilitation and two (2) days coaching training, plus two (2) days instructional leadership training (forthcoming) for HODs</td></tr>
<tr><td>Literacy Instruction Monitoring & Support</td><td colspan="2">One to two visits per month per Grade 1 and 2 classes by coach
One visit per quarter per Grade 3 class by coach
One visit per month per class by FP HOD</td></tr>
<tr><td>Library Management Training</td><td colspan="2">Eight (4, 2, 2) days for teacher-librarians and FP and Intermediate/Senior (INTERSEN) HODs</td></tr>
<tr><td>Library Monitoring & Support</td><td colspan="2">One visit per month</td></tr>
<tr><td>Reading Activities Training</td><td colspan="2">Six (4, 2) days for teacher-librarians and HODs
Six (4, 2) days of school-based training for teachers</td></tr>
<tr><td>Reading Activities Monitoring & Support</td><td colspan="2">Two visits per school per month for library period observation by coach</td></tr>
<tr><td rowspan="2">Resources</td><td rowspan="2">Materials in Classroom and Library</td><td colspan="2">One set of Learner Books per grade per learner
One set of Teacher Guides per grade per teacher
One set of letter cards per grade per class</td></tr>
<tr><td colspan="2">Five library books x enrolment over 3 years
One set of library furniture, mats and checkout system in Year 1
Library stationery each year</td></tr>
</table>

1 A language or languages learned at home, also known as mother-tongue language or first language.

Community & Government Partnerships	Family & Community Engagement	One Community orientation meeting per school One day Principal Orientation meeting per year
		Two parent meetings per year Two literacy events per year
	Government engagement	One progress update meeting per quarter per district and province
Monitoring and Assessment	Classroom instruction	Learner tracking assessment – once per year, per grade Monitoring during support visits (see above)
	Library	Checkout data collected - once per month Monitoring during support visits (see above)

Source: Unpublished Room to Read annual planning criteria

**Reads as: teachers receive instruction training over three years with 19 days total for Grade 1 teachers, broken down into ten days in year 1, six days in year 2, and four days in year 3.*

1.1.1 Resources

The Literacy Programme provides resources for schools as described in Table 2. The content of the instruction component is research-based and uses a structured pedagogy approach[2] (Kim & Davidson 2019; Learning Point Associates 2004; Dornbrack & Kazungu 2022). Each lesson is designed to cover the five components of reading[3], is aligned to the Curriculum Assessment Policy Statements (CAPS), and is designed to take 45 minutes. The course contains three texts per week for 40 weeks of teaching.

The library, usually situated in a spare classroom or renovated storeroom, is provided with furniture and stationery. Library books, both in the home languages of the school and in English, are provided periodically over the three years with a goal of at least five books per learner.

1.1.2 Teacher support

Implementation of instruction is rolled out in Grades 1 and 2 in subsequent years. As noted in Table 2, training and coaching inputs are highest in a teacher's first year and decrease after that. Coaches are a largely-permanent corps of experienced teachers trained to coach and facilitate workshops following the Literacy Programme model. Training workshops, which are spread through the first three terms of each year, consist of some theory, orientation to the materials, modelling of methods and lessons, and practice.

Foundation Phase HODs attend the training of teachers and receive coaching and facilitation skills training (see Table 2). Over the three years, the coach uses the gradual release approach, with the coach doing increasingly less and the HOD more

2 See section 3.5-3.8 for more detail.

3 Research has identified five key components of reading: phonological awareness, phonics, fluency, vocabulary, and comprehension (Learning Point Associates 2004).

over time, to transfer responsibility for coaching from the coach to the HOD during joint observations and feedback sessions (Reid & Kleinhenz 2015).
Before the opening of the library in Year 1, a volunteer teacher-librarian, the Foundation Phase (FP) HOD, and the Intermediate Senior Phase (INTERSEN) HOD attend training on library-management systems and reading activities. The library-management system covers library set-up and maintenance, a book checkout system, and creating a child-friendly space. The reading activities training focuses on library rules, learners' use of books, and four reading activities: reading aloud, shared reading, paired reading, and independent reading. The coach then assists these three educators to do school-based training on the reading activities to the rest of the teachers and provides coaching to teacher-librarians and teachers. Table 2 specifies the number of days for each library training over the three years.

1.1.3 Monitoring and assessment

Monitoring and assessment are important aspects of the classroom instruction programme (see Table 2). Assessment by the teacher is built into the lessons once a week and there is a formative learner-tracker assessment conducted by teachers at least once a year, with analysis and reporting to parents. Coaches monitor programme implementation using observation forms that incorporate key aspects of implementation and guidance for providing feedback. For instruction, monitoring focuses on teachers using the instructional routine, mastering basic skills (e.g. correct letter sounds), and mastering advanced skills (e.g. probing questions). Library-management monitoring focuses on library room set-up and condition, a functional checkout system, and library period scheduling. Coaches observe library periods and provide feedback to teachers using a reading-activity observation form for each activity.

1.1.4 Family, community, and government engagement

Room to Read engages with family, community, and the government in multiple ways. The school community and the district are actively engaged in the selection of schools. Schools set up a literacy committee to promote support for the library and reading. A one-day workshop for principals, in addition to regular meetings at the school, is facilitated each year to outline the programme for the year, give progress updates and find solutions to common problems. Coaches support each school to host two literacy events per year and attend parent meetings to advocate for support of reading at home. Quarterly reports are provided to the district.

1.2 Research question

This chapter seeks to answer the following question: what are the key learnings from the Literacy Programme intervention, having spent over a decade in addressing the skills and habit of reading in South Africa? The authors examined reports and other organisation documents, conducted two focus groups with community-based staff, and drew from their experiences in their collective 28 years at Room to Read.

2 Literature review

As described above, the main inputs of the Literacy Programme are teaching and learning resources, teacher support, monitoring and assessment, and government, family, and community engagement. This section will review the research behind these elements.

2.1 Resources

Research indicates learners need access to multiple types of appropriate learning materials. Textbooks in the appropriate languages and level of difficulty can improve learning and are considered cost-effective (Abeberese et al. 2011; UNESCO 2016). When teachers are supplied with scripted teacher's guides, this can offset challenges some teachers may have due to poor content knowledge (Taylor & Vingevold 1999; Veriava 2015; Kim et al. 2016). To develop the skills and habit of reading, children need a wide variety of books to read in a language they speak. Creating an environment that promotes and enables reading in school leads to improvements in reading skills and motivation (Abeberese et al. 2011; Blessing 2020). Despite the efforts of both the South African government and NGOs to create a more equitable supply of books, there are still limited materials in African languages and large differences in the amount of materials between languages (Department of Basic Education 2019).

2.2 Teacher support

Many pre-service teacher training programmes do not adequately provide training on early grade reading, particularly in African languages. In a study of 16 South African universities, only seven offered courses that explicitly focused on literacy and only three of the African languages were represented (Reed 2019). It was only in 2019 that a concerted effort was made to document the different teaching needs of indigenous languages through the development of the National Framework for the Teaching of Reading in African Languages (Department of Basic Education 2019). The national Primary Teacher Education Project, created to address teacher education quality, has developed draft standards for teacher education in language and literacy (PrimTEd Literacy Working Group 2020).

While pre-service training is vital, there is evidence that continuous in-service professional development may have an even greater impact on teacher capacity and ability (Hartwell 2013; Bean 2014; EFA Global Monitoring Report team 2015). Professional development can be provided through ongoing training, modelling and feedback by coaches and local professional learning communities (McLaughlin & Talbert 2006; Bean 2014; Voelkel & Chrispeels 2017). The Early Grade Reading Study (EGRS), a randomised control trial of three interventions in South Africa, found that an intervention with teacher training and coaching support outperformed an intervention with only teacher training and an intervention with parents (Department of Basic Education 2017). However, teachers in some South African studies found their in-service training options to be inadequate to their needs and viewed their attendance as related to accountability and not to change classroom behaviours (Mampane 2018; Mutereko 2019).

2.3 Monitoring and assessment

The main purpose of formative assessment is for teachers to understand their learners' progress and to use data to modify their instruction to better meet the learners' needs (Crouch 2011; Florez & Sammons 2013). A study in Mpumalanga found that teachers did not have a good understanding of formative assessment and that without quality feedback for the learners, these methods of assessment are of little use (Mkhwanazi 2018).

Making changes based on needs observed in the school is another way to improve outcomes. Monitoring data, in the classroom or library, should be collected, analysed against set standards, and used for programme improvement (Keller 2018). Monitors should focus on actionable indicators that implementers can control (Gibson n. d.).

2.4 Government, family, and community engagement

Schools operate within a socio-cultural community as well as a school community, which includes the district and provincial departments of education. Stakeholder involvement *from the outset* matters for a project's long-term success (World Bank, Operations Evaluation Department 1991). Project sustainability relies on having a variety of stakeholders, including parents and other community members, who understand and commit to the project's purpose, activities, outputs, and goals. (National Clearinghouse for Comprehensive School Reform [NCCSR] 2002).

A child's motivation to read is highly influenced by the home literacy environment and parental support (UNESCO 2011). A study funded by the Bill and Melinda Gates Foundation found that children whose parents read with them or who read regularly on their own are 50 per cent more likely to read for leisure at home and 35 per cent more likely to find reading interesting, compared to their peers with less-involved parents (Juarez & Associates 2015).

3 Reflections and lessons learned

Room to Read's literacy programming has evolved over time. There have been successes, such as a 2017 evaluation using the Early Grade Reading Assessment (EGRA), where South African Grade 2 learners in project schools read with significantly more fluency and comprehension and obtained significantly fewer zero scores than Grade 2 learners in comparison schools (Rigole & Phaweni 2018). There have also been many challenges, some of which influenced big changes in programming, as described in this section. There continues to be work to be done. The authors reflected on their experiences, reviewed notes from monthly community-based staff meetings, and conducted two focus groups to identify 10 lessons learned that have shaped the programme.

3.1 Modules that guide programme implementation lend clarity

The early iterations of the Literacy Programme differed between countries and made it difficult for staff to support the programmes. This led to a programme-wide shift across all countries supported by Room to Read to strengthen the quality of the Literacy Programme on all levels, from materials to routines to trainings, by introducing programme-implementation modules. Standardising how instruction lessons, library periods, and professional development were developed and conducted introduced an element of quality control across countries and, in-country, across provinces and districts. These modules guided the creation of the current Sepedi Literacy Programme in 2015 and the isiZulu programme in 2019. The country office took local factors into account, such as aligning graphemes by term and writing tasks to CAPS, to create a programme specific to the South African context but aligned with the modules. Programme developers return to these modules each time they add a new language. This ensures that each language programme is crafted with the same quality practices and makes the intervention easier to scale.

3.2 Building support with communities requires planning and persistence

Learners whose parents and communities are involved, value reading and read more (Wigfield et al. 2016), but parents and community members often work long hours and may not be interested in engaging in literacy activities. In alignment with this research, Room to Read has found that engaging local government, communities, and families strengthens the programme, but it can be difficult to get people to invest the time and effort. Building support takes time and careful planning.

For example, most schools and parents participate in fixing up the library space and purchasing library furniture, but some schools struggle to get parent participation. In response, coaches consistently engage with the community in activities such as promoting literacy at the regular school parents' meetings. Parents are more likely to attend a meeting if there is a library tour led by learners or a showcase of learner reading skills. Parent participation has increased to 50-80 per cent. In addition, a WhatsApp discussion group with School Governing Board chairpersons was particularly successful in engaging community members in supporting the programme.

School communities are engaged through representation on the school's literacy committee, attendance at literacy events and recruitment of library volunteers. Room to Read has encouraged coaches to engage directly with the literacy committee (e.g. by sharing guidance on building home support for reading). Community members need the most help with the sustainability plan and coaches often assist with identifying possible funders for the library (Legodi et al., personal communication, 2022).

3.3 Access and reading motivation contribute to a habit of reading

Learners cannot build a habit of reading if they do not have access to books. An external evaluation completed in 2015 showed that Room to Read's libraries led to an increase in the percentage of children reading for enjoyment at home, compared to control schools and compared to schools pre-intervention (Juarez & Associates 2015).

However, it is challenging to encourage reading among children not already immersed in a strong culture of literacy (Evans et al. 2010). In 2015, a multinational study showed that children were more likely to read when they attended frequent library periods and had a teacher trained to conduct reading activities (Room to Read 2015). In the early years in South Africa, it was difficult to get buy-in from schools on regular use of the library until the programme began requiring 30-minute library periods. Teacher-librarians volunteered to oversee the management of the library. Some subject teachers objected to conducting the library period, so language teachers took learners to the library. Despite these increases in time spent in the library, in 2010, the average number of books checked out per learner per year in South Africa was only 0.4 books (Room to Read 2013, 2020a).

Literacy coaches used different strategies to encourage checkout, such as changing the checkout system so older learners could help, and requiring schools to source library assistants because assistants facilitated higher checkout and longer library hours (Phaweni & Morgan 2014). When coaches helped principals, teachers, and parents understand the purpose of the library, checkout, and the role of each participant, checkout increased to 10.7 books per learner in 2019.

Increased reading was also noted when the library period focused on reading for enjoyment rather than teaching reading skills (Room to Read, 2015). Therefore, coaching and teacher training at project schools emphasised the importance of allowing children to select their own reading material and of focusing on the joy of reading during library time (Wigfield et al. 2016; Brandt et al. 2021).

3.4 Culturally and linguistically relevant books are levers for supply and demand

Access to stimulating collections of attractive, appropriate, and relevant books is an important determining factor in the achievement of early literacy and a habit of reading. Many low- and middle-income countries experience a significant undersupply of storybooks, particularly in home languages, and, when books are available, there are critical issues including poor quality and low usage (Davidson 2013; Manu et al. 2019; Gray 2021).

The Literacy Programme aims to improve both the supply of and the demand for high-quality children's books that will motivate children to love reading (Pallangyo 2016). Originally, libraries were stocked only with English book donations, but the need for more linguistically and culturally relevant books became evident. Publishing books locally has been a key strategy for addressing this need. At first, Room to Read

South Africa conducted competitions and accepted manuscripts to create books, but the quality was not satisfactory. To improve the quality, the organisation began to train South African writers and artists to produce home-language books for primary school-age children, with a total of 35 authors and 16 illustrators trained since 2015[4]. Writers and illustrators who attended the training developed their skills in areas such as story structure, character studies, and storyboarding (Mositsa & Ziki personal communication, 2022). Room to Read South Africa has published 93 original titles and 414[5] adapted books, representing 11 languages and, in 2020, began uploading books to an online literacy platform.

3.5 Systematic and explicit literacy instruction builds skills

An internal review of schools implementing libraries in 2008 found that the youngest learners were not going to the library because they could not read. This led to the introduction of an instruction programme focused on developing literacy skills among Grade 1 and 2 learners.

While the first instruction programme introduced in South Africa was based on the five components of reading, it was organised around curricular themes and did not follow a systematic progression of skills. In 2014, the organisation took an approach focused on systematic and explicit routines, a scaffolded pedagogy, and making assessment-informed decisions (Kim & Davidson 2019). Literacy staff created guidelines on how to analyse a language to develop a scope and sequence that is systematic, introducing two graphemes per week in a productive order and supporting learners to gradually build on their previous learning (Dornbrack & Kazungu 2022; Kim & Davidson 2019). The South Africa team designed their programme with attention to learning to read in *African* languages, for example, blending graphemes into syllables and syllables into words due to the frequency of long words.

The Teacher's Guide uses a scripted approach, which specifies a method of explicit instruction, by explaining to learners what is expected of them, providing a model, shared practice, and independent exercises and feedback (Learning Point Associates 2004). Community-based staff feel that this scaffolded approach, often referred to as 'I do, we do, you do,' supports learners until they reach mastery of the skill (Banyini et al., personal communication, 2022).

Initially, an implementation challenge was that many teachers would model all the words from activities such as blending, and skip the independent practice or allow learners to chant all together. To address this problem, literacy coaches emphasised the importance of independent practice and regularly reminded teachers to have the learners read quietly at their own pace so that chanting did not prop up the weaker readers (Legodi et al. personal communication, 2022). Teachers have improved, but this scenario is repeated when the programme enters a new school.

4 Data before 2015 is unavailable.

5 These 414 adapted titles include multiple translations of a single unique title into different languages.

3.6 Content and design matter for learner books

When improving learner books in 2015, the organisation made changes to both design and content. Room to Read South Africa's initial classroom materials were developed according to national guidelines focused on themes and subject integration, using consumable workbooks and one set of phonic reading cards with basic descriptive texts. In a two-year (2013-14) impact evaluation of the Sepedi language Literacy Programme in Limpopo and Mpumalanga, both project and comparison learners made greater gains in letter-sounding than fluency and comprehension, indicating a need to focus more on connected text (Room to Read 2014).

In the updated, colour, non-consumable learner books, activities are repeated in a predictable format and the content becomes more complex as the year progresses. Best practices regarding font size and white space make the books user-friendly for both learner and teacher (blueTree Group 2014). As with the earlier workbooks, each learner receives a learner book. The most significant change from the original programme to the improved programme is the addition of decodable stories. Decodable stories, written with only the graphemes children have already learned, provide children with a real reading experience that is at an appropriate level for the very earliest beginning readers (Davidson 2013; Dornbrack & Kazungu 2022). Drawing from organisational experience with children's book publishing, teams wrote these short stories following learning regarding creating quality stories. Learner books contain syllable and word reading for reading skills practice and three decodable stories per week (120 in all) that can be used for shared, group-guided, and independent reading.

3.7 A consistent routine promotes mastery

In 2014, Room to Read undertook to standardise the quality of instruction across countries and languages. A decision was made to streamline lessons to include only activities that programme experts felt contributed most to learner achievement, following the research and the Room to Read's years of experience in schools. The organisation created a simple routine for each new grapheme introduced, which experts believed would produce results for learners. These limited activities covered the five components of reading plus writing each day with explicit instructions contained in a Teacher's Guide. The South African team divided the lesson into ten consistent steps to simplify the process for teachers while still providing guidance and content.

Some education officials expressed concern that the new programme would not engage learner attention and would take away teachers' creativity. Teachers in South Africa reported liking the repetition of the programme because it affords learners adequate opportunities to grasp the new skills and helps slow learners to catch up (Banyini et al., 2022). Working with a consistent routine allowed teachers to maximise teaching time in the 45-minute period allotted for the Literacy Programme. At first teachers struggled with finishing the lesson in 45 minutes, but as they became more confident and practiced, they were able to fit the lesson into the planned time (Legodi et al. 2022).

3.8 Data-based decision making brings improvements and challenges

Collecting data in the classroom is a familiar task, but it is only effective in improving learning when teachers use the data to make changes to their instruction. In 2013, Room to Read introduced materials and training for a formal ‘learner tracking’ assessment (Florez & Sammons 2013; Room to Read 2016). This learner tracking has never functioned as intended in South Africa. The one-on-one assessment for all learners in Grades 1 and 2 was designed to be administered twice per year, covering material learned so far. Teachers and coaches struggle with the amount of time required for administration and therefore implement it once per year.

Teachers and coaches do some simple categorisation of learner tracking results. Studies have found that teachers often lack the skills to connect data to instructional decisions (Crouch 2011). As such, suggested ‘actions for learning’ are provided directly on the record sheet for various levels of achievement. The actions are designed to be simple, low-resource, and easy to implement in a variety of classrooms (Room to Read 2016). Some teachers find these suggestions helpful to support struggling readers, but most teachers do not use the data at all (Banyini et al. 2022; Legodi et al. 2022).

Consistent monitoring of library functions allows teacher-librarians to improve the performance of the library. In 2012, a form was introduced to standardise support for library management. Coaches began to use this tool once a month during their visit with the teacher-librarian to guide feedback. Bi-annually, data on this form are collected and libraries are categorised into Developing, Functional, and Highly Functional[6]. Developing libraries then receive more coach visits. Based on staff feedback, the form was revised in 2015 to be more specific and concrete, which made it easier for coaches to complete.

3.9 Essential elements help to change teacher behaviour

Fostering behaviour change can be difficult. Using a classroom or library observation form, coaches can track teachers’ adoption of the desired behaviours for improved literacy and library implementation over time. In 2019, teachers who were rated as mastering advanced skills in literacy instruction rose from 4 per cent at the beginning of the year to 47 per cent at the end of the year. Libraries that were functioning or highly functioning rose from 28 per cent to 88 per cent (Room to Read 2020b). This behaviour change comes about through training and coaching.

Coaches reported that one of the most important aspects of creating teacher buy-in is helping them to understand the reasons behind the content and pedagogy (Banyini et al. 2022). The globally developed training sessions include interactive discussions and focus on practice, with lots of time provided for demonstrations, practice, and feedback. However, this has been difficult to implement in the time allowed. Facilitators

6 Room to Read monitors the smooth and efficient operation of the library using a structured observation form, whose results divide libraries into Developing, Functioning, and Highly Functioning. These results impact the number of support visits for libraries.

have adopted a lecture-based approach for some portions and shortened the practice time to enable them to get through the material.

Teachers in the Literacy Programme receive regular follow-up support from their literacy coach through classroom observations and feedback, which are particularly important given the difficulty in managing practice time during workshops. Coaching support varies by year (see Table 2), but research has found coaching to be essential to help teachers to adopt new behaviours (Reid & Kleinhenz 2015; Smith & Gillespie, 2007). When Room to Read began employing coaches, the quality of feedback depended on the expertise of the coach in judging the greatest needs. In 2016, coaches began using detailed observation forms with clear criteria and guidance for the coaches to identify and structure their feedback. Coaches were trained to provide one or two pieces of actionable feedback, but this had to be reinforced by supervisors to change the habit of listing all the teacher's mistakes.

3.10 Sustainability must be planned

An intervention may be successful, but it requires more than that to sustain change. Keys to behaviour change include teachers understanding the reasoning behind the changes, ongoing in-school support, and support from government and school leadership (NCCSR 2002). NGOs must take care to enable school leadership and local government to take ownership of the intervention (Mkhwanazi et al. 2018). As described in section 3.9, Room to Read focuses on helping principals, HODs, and teachers to understand the reasons for the components of the Literacy Programme and to plan for sustainability from the beginning.

Christie and Monyokolo (2018) found that the involvement of school leadership and the role that the school management team could play in sustaining change was crucially important. This influenced Room to Read South Africa's design for a system of support for Foundation Phase HODs to familiarise them with their mentoring role and to build their skills in coaching their fellow teachers in instruction, with the intention that they take over when programme support phases out. In 2020, coaches began working with Intermediate/Senior Phase HODs to support volunteer teacher-librarians and teachers in leading reading activities. The biggest challenge for HODs, with a full teaching load and administrative tasks, is finding time to support teachers. However, HODs appreciate the skills they have gained and the impact of their work through the programme on teachers and learners, including from beyond the Foundation Phase. They also plan and facilitate successful professional learning communities with minimal support (Banyini et al. 2022).

4 Way forward and conclusion

Through these lessons, Room to Read has evolved a literacy programme that has improved learner reading outcomes, increased book usage, and changed teacher behaviour. This forms a good basis for implementing at scale and pursuing systems change. In this section, the authors set out some of the challenges and opportunities for large-scale change and future work.

4.1 Implementing at scale

Scaling up a programme to reach a greater number of people requires an understanding of what makes an intervention successful in a particular context and the preparedness and willingness of the system to receive and integrate the intervention (Cooley et al. 2021). Room to Read India implemented a three-tiered plan to scale the programme and found that a programme with fewer inputs had a similar impact on fluency and comprehension as one with more inputs (Cooper & Joddar 2017; Joddar 2018). They attributed this to the government taking the lead in the intervention. Following this research, Room to Read South Africa has begun to implement a similar approach to scaling its literacy model.

Demonstration schools, the first tier of the plan and described extensively above in Table 2 and Section 1.1, serve as proof-of-concept to generate buy-in in a particular district and province. Room to Read supports the entire Literacy Programme financially. This allows districts to see the benefits of the programme in their context before committing to adding schools. To date, the organisation has implemented 355 comprehensive Literacy Programme Demonstration schools since 2014, consisting of classroom instruction and libraries in eight districts in four provinces.

Collaboration schools, the second tier, are a collaborative effort to support the handover of the programme to district government. Room to Read began to implement Collaboration schools in 2019 in three districts in Limpopo and Mpumalanga, where demonstration schools had been implemented and local government was supportive of the programme. Schools were selected based on their interest in the programme.

In collaboration schools, the responsibility for professional development and support, resources, community and government events, and monitoring and assessment (see Table 2) are shared between the district, province, and Room to Read. There has been enthusiasm from school communities and districts in Limpopo and Mpumalanga for the idea of expanding the programme, but thus far, they have struggled in implementation. The organisation has conducted training sessions, which district officials and educators have attended, but neither the province nor the districts have been able to commit sufficient funds to take on the responsibility for providing material resources. Some school communities have provided furniture, but only about half of 2021 collaboration schools have libraries equipped with furniture. The Literacy Programme provided 2021 schools with some surplus books, but that supply is exhausted. Schools in 2022 have printed out black-and-white copies of the lessons for learners to share, which makes independent practice more difficult to achieve.

After collaboration schools have been successfully implemented, expansion schools, the third tier in the scale-up plan, would be implemented district-wide, with the district or province covering costs and implementation and Room to Read providing technical assistance. However, given the challenges in the collaboration phase, more discussion with each province is needed to determine how the programme can meet the province's needs.

4.2 Influencing system change

While the Literacy Programme has helped learners to improve their skills and habit of reading, there is room for improvement. The organisation conjectures that a deeper and wider partnership with the Department of Basic Education may be the key to widespread, lasting change (Hartwell 2013). To support system change, Room to Read is taking several steps that could pave the way.

4.2.1 Building evidence for in-school coaching and training

As described above, the Literacy Programme has begun to train HODs to support the teaching of literacy and the use of libraries to foster sustainability of the gains achieved. The organisation planned an evaluation that was delayed by the Covid-19 pandemic. It is hoped the evidence behind this approach will grow, providing an evidence-based option for coaching using personnel resources that are already in place.

4.2.2 Strengthening children's book publishing

The organisation is working to strengthen children's book publishing to increase access to quality books. Through the World Bank-funded Results in Education for All Children (REACH) project, Room to Read South Africa published guidelines for creating quality storybooks that were endorsed by the minister of education[7]. The organisation plans to partner with local publishers to either license Room to Read's titles or co-create titles to increase available books. In 2022, Room to Read South Africa will explore the options that work best in this context.

4.2.3 Advocating for school or classroom libraries

There are many NGOs and government agencies working to help children learn to read. At various forums to discuss literacy in South Africa, libraries are seldom mentioned. Room to Read South Africa will use its inaugural conference in July 2022 to raise awareness and discuss the establishment of functional libraries across the country with reading-focused NGOs and Department of Education officials, drawing on the organisation's experience with national library partnerships to scale school libraries in Cambodia and Vietnam.

4.3 Combining science and magic

Children can only develop a habit of reading if they have reading skills, and children's skills improve through a habit of reading. These two goals of literacy—reading skills and reading habit—complement each other to help children become lifelong, independent readers. The coming together of quality literacy instruction, child-friendly library programmes, and engaging children's books to nurture these goals is one that happened through incremental learning for Room to Read. This chapter has shared 10 lessons learned, as well as noting where more must be done regarding implementing literacy initiatives that marry skills and habit to motivate children to read frequently, willingly, and with enjoyment. Room to Read's Literacy Programme brings together the science and the magic of reading by not only teaching children how to read, but also helping them fall in love with reading.

7 The guidelines are available at www.literacycloud.org under Resources.

References

Abeberese, A.B., Kumler, T.J. & Linden, L.L. (2011). *Improving Reading Skills by Encouraging Children to Read in School: A Randomized Evaluation of the Sa Aklat Sisikat Reading Program in the Philippines (NBER working paper series, no. w17185)*. Cambridge, Mass: National Bureau of Economic Research.

Banyini, A., Shongwe, N., Nonyane, K., Mashaba, B., Mdunge, S., Nkosi, M., Tladi, K., Mabuza, L., Vilakazi, M., Raboshakga, E., Timakwe, N., Ncube, P. (2022). Personal communication.

Bean, R. (2014). *The power of coaching: Improving early grade reading instruction in developing countries*. Washington, D.C.: USAID.

Blessing, C. (2020). Designing Child-Friendly School Libraries and Media Centres for Sustainable Reading Development amongst Nigerian children. Library Philosophy and Practice, (Online). Available at: https://digitalcommons.unl.edu/libphilprac/3913/ (Accessed 13 November 2021).

blueTree Group (2014). *Best Practices for Developing Supplementary Reading Materials. Washington*, D.C.: USAID.

Brandt, L., Sharp, A.C. & Gardner, D.S. (2021). Examination of Teacher Practices on Student Motivation for Reading. *The Reading Teacher*, 74(6), 723–731.

Christie, P. & Monyokolo, M. (2018). *Learning about Sustainable Change in Education in South Africa: The Jika iMfundo campaign 2015-2017*. Johannesburg: Saide.

Cooley, L., Seghers, M. & Perlman, R. (2021). Planning for Scale: The Education Scalability Checklist, (Brookings blog). (Online). Available at: https://www.brookings.edu/blog/education-plus-development/2021/02/24/planning-for-scale-the-education-scalability-checklist/ (Accessed 21 October 2021).

Cooper, P. & Joddar, P. (2017). *India Literacy Program: Demonstration School Endline Evaluation Report*. San Francisco: Room to Read.

Crouch, L. (2011). Motivating Early Grade Instruction and Learning: Institutional Issues. In A. Gove & A. Wetterberg (Eds.) *The Early Grade Reading Assessment: Applications and Interventions to Improve Basic Literacy.* (1st edn.). Research Triangle Park, NC: RTI Press, 227–250.

Davidson, M. (2013). *Books that Children Can Read: Decodable Books and Book Leveling*. Washington, D.C.: USAID.

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

Department of Basic Education. (2019). *National Framework for the Teaching of Reading in African Languages in the Foundation Phase.* Pretoria: Department of Basic Education.

Dornbrack, J. & Kazungu, T.C. (2022). Decodable Stories and Their Role in Reading Skills Development in South Africa. In Spaull, N. & Pretorius, E. (Eds.) *Early Grade Reading in South Africa*. Cape Town: Oxford University Press.

EFA Global Monitoring Report team. (2015). Investing in teachers is investing in learning: A prerequisite for the transformative power of education: background paper for the Oslo Summit on Education for Development. Oslo, Norway: UNESCO.

Evans, M.D.R., Kelley, J., Sikora, J. & Treiman, D. J. (2010). Family scholarly culture and educational success: Books and schooling in 27 nations. *Research in Social Stratification and Mobility*, 28(2), 171–197.

Florez, M.T. & Sammons, P. (2013). *Assessment for learning: effects and impact*. Oxford, UK: Oxford University Department of Education and CfBT Education Trust.

Gibson, M. (no date). Implementation monitoring, The Abdul Latif Jameel Poverty Action Lab (J-PAL). (Online). Available at: https://www.povertyactionlab.org/resource/implementation-monitoring (Accessed 17 March 2022).

Gray, L.C. (2021). Mirrors, Models and Mother Tongues: Why cultural relevance matters. (Working Paper 2/21). Brisbane, Australia: Library for All. Available at: https://libraryforall.org/wp-content/uploads/2021/07/20210702-WPS-WP2_21-1.pdf (Accessed 5 January 2022).

Hartwell, A. (2013). Policy matters: Early grade reading policy toolbox. Internal University of Massachusetts report. Unpublished.

Joddar, P. (2018). Impact Evaluation of the Literacy Program-Partnership Approach under Scaling up Early Reading Intervention (SERI) funded by USAID: 2018 *Endline Evaluation Report for Chhattisgarh and Uttarakhand*. San Francisco: Room to Read.

Juarez & Associates. (2015). School Libraries Cross-National Evaluation. San Francisco, CA: Room to Read. (Online). Available at: https://www.roomtoread.org/media/ypqhmqgt/school_libraries_cross-national_summary.pdf (Accessed 3 November 2021).

Keller, C. (2018). AASL Standards for Supporting Library Evaluation: Linking to Building, District, and State Levels. *Teacher Librarian*, 45(5), 17–21.

Kim, Y.-S.G., Boyle, H.N., Zuilkowski, S.S. & Nakamura, P. (2016). *Landscape Report on Early Grade Literacy*. Washington, D.C.: USAID.

Kim, Y.-S.G. & Davidson, M. (2019). *Promoting Successful Literacy Acquisition through Structured Pedagogy*. Washington, D.C.: USAID.

Learning Point Associates. (2004). *A Closer Look at the Five Essential Components of Effective Reading Instruction: A Review of Scientifically based Reading Research for Teachers*. Illinois: Learning Point Associates.

Legodi, V., Ncube, P., Nkambule, L., Banyini, A., Nonyane, K., Ngwane, C., Sikhipa, G., Sebashe, Z., Kgoete, M., Furumele, V. (2022). Personal communication.

Mampane, T.J. (2018). *School Heads of Department's Role in Ensuring Teacher Professional Development in Mathematics: The South African Context*. Sofia, Bulgaria: Bulgarian Comparative Education Society (BCES Conference Books). Available at: http://eric.ed.gov/?id=ED586164 (Accessed 14 October 2019).

Manu, A., Ewerling, F., Barros, A.J. & Victoria, C.G. (2019). Association between availability of children's book and the literacy-numeracy skills of children aged 36 to 59 months: secondary analysis of the UNICEF Multiple-Indicator Cluster Surveys covering 35 countries. *Journal of Global Health*. 9(1), 010403.

McLaughlin, M.W. & Talbert, J.E. (2006). *Building school-based teacher learning communities: Professional strategies to improve student achievement*. New York: Teachers College Press.

Mkhwanazi, T., Ndlovu, Z. & Bansilal, S. (2018). Exploring mathematics teachers' usage of the curriculum planner and tracker in secondary schools in King Cetshwayo and Pinetown districts. In Christie, P. & Monyokolo, M. (Eds.). *Learning about Sustainable Change in Education in South Africa: The Jika iMfundo campaign 2015-2017*. Johannesburg: Saide, 195–224.

Mullis, I.V.S., Martin, M.O., Foy, P. & Drucker, K.T. (Eds.). (2007). *PIRLS 2006 International Results in Reading*. Massachusetts: Boston College TIMSS & PIRLS, International Study Center.

Mullis I.V.S., Martin, M.O., Foy, P. & Hooper, M. (2017). PIRLS 2016 International Results in Reading. Massachusetts: Boston College TIMSS & PIRLS International Study Center. Available at: http://timssandpirls.bc.edu/pirls2016/international-results/. (Accessed 6 January 2021).

Mutereko, S. (2019). Accountability through Continuous Professional Development: Perceptions of Educators in Umgungundlovu District, South Africa. *International Journal of Business and Management Studies*, 11(1), 85–100.

National Clearinghouse for Comprehensive School Reform (NCCSR). (2002). Sustaining Your Reform: Five lessons from research. *Benchmarks*, 3(3).

Pallangyo, A. (2016). *Developing Quality Books for Every Child*. USAID: Global Reading Network.

Phaweni, K. & Morgan, C. (2014). Book Checkout Analysis Study Report. Room to Read: Unpublished.

PrimTEd Literacy Working Group. (2020). *Knowledge and Practice Standards for Primary Teacher Education Graduates: Language and literacy*. Pretoria: Department of Higher Education and Training.

Reed, Y. (2019). *Foundation and Intermediate Phase B.Ed. programmes at selected South African universities: languages and literacies components*. Pretoria: Department of Higher Education and Training.

Reid, K. & Kleinhenz, E. (2015). *Supporting Teacher Professional Development: Literature Review*. Canberra, Australia: Department of Foreign Affairs and Trade.

Rigole, A. & Phaweni, K. (2018). *Literacy Program: South Africa (Sepedi)*. Pretoria, South Africa: Room to Read.

Room to Read. (2013). *Global Monitoring Report 2012*. San Francisco, CA.

Room to Read. (2014). Sepedi End-of-Grade 2 EGRA Report. (Online). Available at https://www.roomtoread.org/media/mt5jjpzc/reading-skills-evaluation_za_sepedi_2014_endline_technical-report.pdf (Accessed 15 June 2020).

Room to Read. (2015). *Reading Promotion Study Report*. San Francisco, CA.

Room to Read. (2016). Student Tracking Step-by-Step. Room to Read: Unpublished.

Room to Read. (2020a). *2019 Global Indicators Report*. San Francisco, CA: Room to Read.

Room to Read. (2020b). Program Implementation Monitoring 2019. Room to Read: Unpublished.

Room to Read. (2021). [Number of new learners in 2021]. Unpublished raw data.

Room to Read. (2022). *2020 Global Indicators Report*. San Francisco, CA: Room to Read.

Smith, C. & Gillespie, M. (2007). Research on Professional Development and Teacher Change: Implications for Adult Basic Education. In *Review of Adult Learning and Literacy, Vol. 7*, Connecting Research, Policy, and Practice.

Taylor, N. & Vingevold, P. (1999). *Getting Learning Right*. Johannesburg: Joint Education Trust.

UNESCO. (2011). *Creating and Sustaining Literate Environments*. Bangkok, Thailand: UNESCO.

UNESCO. (2016). *Global Education Monitoring Report*. Policy Paper 23. Paris: UNESCO.

Veriava, F. (2015). Why textbooks are a crucial part of every child's learning journey. *The Conversation Africa*, 8 November.

Voelkel, R.H. and Chrispeels, J.H. (2017). Within-School Differences in Professional Learning Community Effectiveness: Implications for Leadership. *Journal of School Leadership*, 27(3), 424–451.

Wigfield, A., Gladstone, J. & Turci, L. (2016). Beyond Cognition: Reading Motivation and Reading Comprehension. *Child Development Perspectives*, 10(3), 190–195.

World Bank, Operations Evaluation Department (1991). Project Sustainability in *Education*. No. 14. New York: World Bank.

06

Molteno's Vula Bula: Conceptualisation, development, and implementation

JENNY KATZ & SIÂN REES

Abstract

South Africa's national Department of Basic Education considers graded readers to be a core Foundation Phase resource necessary for children to learn how to read. Since 1994, the majority of the few African language readers available in Grade 1-3 classrooms were either translations of English readers, or original texts that followed a whole-language approach. Research shows that a phonics-based approach is the best way to teach reading in languages with consistent, transparent orthographies, yet there were no decodable or basal reading series available in African languages. After three years of schooling, PIRLS 2016 proved that only 22% of our Grade 4 learners could reach the Low International Benchmark (Howie et al. 2017). Reading, like all other honed skills, needs practice as well as the correct tools to support that practice. While there remains a lack of sufficient foundational reading resources in African languages, it is conjectured that the lack of *appropriate* foundational reading resources in African languages has contributed to these poor results.

Molteno's Vula Bula® Foundation Phase graded reading programmes use a predominantly phonics-based approach to develop decodable, basal readers. It is theorised that if phonic knowledge is a key element of learning to read in an opaque orthography such as English, then it is of paramount importance when learning to read in

KEYWORDS
Molteno, Vula Bula, African languages, graded readers, basal readers, decodable texts, phonics, orthography

the transparent orthographies of African languages.

Since 2012, the Vula Bula graded readers and anthologies –available in all nine of South Africa's official African languages – have been incorporated into notable literacy interventions in several different provinces. They have been well received across all the languages and evaluations are beginning to show promising reading outcomes.

1 Introduction

In 2010, the Molteno Institute for Language and Literacies (Molteno) received a grant from the Zenex Foundation to develop a comprehensive Grade 1 isiXhosa literacy programme. Until then, Molteno was best known for its successful African language Breakthrough to Literacy programmes. The components of these programmes, however, were becoming prohibitively expensive to produce, and were still in the process of being restructured to meet the specifications of the new Curriculum Assessment Policy Statements (CAPS). The objective was to produce a new CAPS-aligned authentic isiXhosa literacy programme that would be more economically viable.

Molteno contacted several well-known publishers and undertook a desktop review of all the isiXhosa graded readers that were on offer at the time. In each case, the readers had been translated from original English texts, resulting in complex isiXhosa word structures that were inappropriate for teaching reading in Grade 1 (Katz 2019). Due to the transparency of isiXhosa orthography, it was decided that Molteno would use an incremental, progressive, synthetic phonics-based approach to develop the new graded readers[1]. The same artwork would be reproduced in corresponding Big Books including extended, higher-level versions of each story, in order to expose children to richer vocabulary and more complex sentence structures.

From the outset it was imagined that this new programme might be adapted into other African languages. Hence it was important to find a name that would have multilingual resonance. The programme was named Vula Bula – *vula* meaning 'open' in isiXhosa, isiZulu, isiNdebele, Siswati and Tshivenḓa, and *bula* meaning 'open' in Setswana, Sesotho and Sepedi[2]. The name was symbolic of Molteno's mission to encourage children to open books to read and write, and to open their minds to lifelong learning and knowledge acquisition.

The Vula Bula programme consists of an Alphabet Wall Frieze, Phonics Cards, High-Frequency Wordlists, Vocabulary Posters, Graded Readers (for group guided, paired and independent reading), and Big Books (for shared reading and reading aloud). At this time, the programme is available in all nine official African languages.

Parts of sections 2 and 5 of this chapter, which describe the rationale, approach, and methodology in the development of the Vula Bula readers, are drawn from the masters thesis of one of the authors (Katz, 2019).

1 Synthetic phonics is a method of teaching reading that starts with basic letter sounds, blending and sounding out words, as opposed to analytic phonics which starts at the word level and then deconstructs the word by looking at its constituent parts.

2 The word for 'open' in Xitsonga is *pfula*, however it was deemed too lengthy to name the programme Vula Bula Pfula. Importantly, *vula* means 'say' in Xitsonga, so the name still resonates for those speakers.

2 Literature

2.1 Access to print is needed for reading practice

Children enter school with oral language and some concepts of print. Most are introduced to the rules of written language for the first time. While oral language develops naturally, reading skills must be learned. Donaldson (1978, 97) warns that awareness of a correspondence between written and spoken language "should never be taken for granted" and that "it is essential to make certain that the child understands that the marks on paper are a written version of speech". Children need explicit help if they are to crack the code which, in an alphabetic system, relates letters or small groups of letters to the sounds of spoken language (Oakhill 1993). Extensive reading practice, using appropriate texts, is therefore needed for children to become literate.

Exacerbating the problem is the minimal amount of accessible available reading material in African languages in schools, particularly at the Foundation Phase level, where children need the most practice in order to become successful readers and achieve academically. The home language curriculum (2011) suggests a minimum of 20 readers per year in Grade 1, which is one reader every two weeks. This is perhaps a start, but it is by no means enough to provide the necessary practice in building the reading fluency and automaticity required for comprehension.

For the young reader, the orthographic lexicon (the storage system in the brain consisting of individual word spellings) develops with increasing reading experience dependent on the availability of reading materials (Aro 2004). The paucity of books in African languages, especially early readers, has been a contributing factor to low reading and literacy rates (Kgobe 2012; National Reading Strategy 2008). Even now, the number of available readers in Foundation Phase classrooms remains extremely low, as shown in recent evidence from the Eastern Cape and as reiterated by Ardington and Spaull (this volume).

The 'Matthew effect' of accumulated advantage accurately describes the reading status quo in South Africa's indigenous languages. Poor readers, less skilled at decoding and understanding texts, are less likely to read as much as their more skilled peers. They will therefore not have as much practice in reading, and the literacy competency gap widens (Heath 1982). The amount of text to which learners are exposed is thus a critical factor in vocabulary development and learning to read fluently, and can have profound cognitive consequences (Cunningham & Stanovich 2003).

2.2 The importance of phonics instruction in African languages

Early grade African-language reading schemes available in South Africa have been failing to provide young children with enough necessary and appropriate practice required to facilitate home-language literacy acquisition (NEEDU 2012). Before Vula Bula, texts had been developed with little appreciation for the agglutinative[3] nature of

3 Agglutinative languages "create words by affixing morphemes to a root morpheme" (Diemer et al. 2015, 328).

African languages and their consistent orthographies. In transparent orthographies, reading is typically taught using purely phonics-based approaches focusing on grapheme-phoneme correspondences (Aro & Wimmer 2003). Texts made up of long words containing multiple complex phonic structures provide inappropriate reading practice for beginner readers, who should receive literacy instruction using easily decodable texts at their instructional reading level (Fountas & Pinnell 1996). From a diagnostic perspective, a text in which a learner can read and comprehend 90% of the words easily is considered to be at that reader's instructional reading level; more difficult text is considered to be at the reader's 'frustration' level and will require additional teacher support (Clay 1991).

The act of reading requires a combination of complex cognitive processes (Wolf 2008; Janks 2011), including the accurate conversion of orthographic to phonological representations (Snow et al. 2005). The fight over best practice in teaching reading in English – the phonics approach versus the whole-language approach – has been belligerent enough to be referred to as "the reading wars" (Pearson 2004). An extensive meta-analysis undertaken by the National Reading Panel in the USA (National Institute of Child Health & Human Development 2000) showed that systematic phonics instruction[4] enhances children's success in learning to read and is significantly more effective than instruction that teaches little or no phonics. First graders who were taught phonics systematically were better able to decode and spell, and also showed significant improvement in their ability to comprehend text. Phonological awareness and knowledge of phonics is critical for providing the requisite decoding skills that are automatically used, even during fluent adult reading (Pelli & Tillman 2007).

Letter-sound correspondence and phonics play a key role in the earliest stages of literacy development and instruction in the Foundation Phase (Dixon et al. 2002). The speed at which children learn to read corresponds approximately to the orthographic complexity of the language that they speak (McDougall et al. 2010). The orthographic depth hypothesis (Frost et al. 1987; Katz & Frost 1992, cited in Ziegler & Goswami 2006, 433) proposes that for languages with transparent orthographies, readers rely more on the phonological or non-lexical route instead of the orthographic or lexical (whole-word recognition) route because the mapping between letters and sounds is mostly explicit. The psycholinguistic 'grain-size' theory of reading (Ziegler & Goswami 2006) describes how children make connections between phonological units or grains of language such as letter sounds, syllables and rimes[5], and the written representation of that language. The authors contend that small grain-size teaching works well in alphabetic languages with consistent letter-sound correspondences. Though syllables are prevalent, the smallest grain-size in African languages is the phonic grapheme. Phonic (letter-sound) recognition is therefore a critical technical skill for fluent decoding in African languages. Text development should therefore show cognisance of the grain-size that readers need to master.

In addition, lexical processing is a challenge for speakers of the Nguni languages due to

4 Systematic phonics instruction is when all the major grapheme-phoneme correspondences are taught in a clearly defined sequence.

5 In English, words can be divided into onsets and rimes. The rime is the word ending or word family. For example in the words 'back', 'quack' and 'shack', the rime is -*ack*.

their conjunctive[6] nature, which impacts word length. Due to their polysyllabic nature, syllabic segmentation is a central part of early reading and spelling instruction (Aro 2004). Reading stamina needs to be gradually developed and trained through sufficient practice in syllabification. Syllables are easily identifiable because they almost always end with a vowel. From the outset, learners should be taught the strategy of quickly recognising vowels in words, and being able to break down words into syllables (manageable chunks). For example, the following words can either be divided into single phonemes or segmented using (vowel)-consonant-vowel consonant-vowel decoding patterns:

- isiXhosa: *ufudo* (tortoise) = *u-f-u-d-o* (phonic) or *u-fu-do* (syllabic)
- isiZulu: *ubisi* (milk) = *u-b-i-s-i* (phonic) or *u-bi-si* (syllabic)
- Setswana: *jeresi* (jersey) = *j-e-r-e-s-i* (phonic) or *je-re-si* (syllabic)
- Tshivenḓa: *ṋamusi* (today) = *ṋ-a-m-u-s-i* (phonic) or *ṋa-mu-si* (syllabic).

If letter sounds are consistent in African languages, one expects less difficulty in learning to read in these orthographies. The challenge lies in the complex consonant combinations or clusters found in the onsets of syllables, and in that there are often multiple complex consonant clusters in words.

Nguni languages boast a prevalence of multiple morphemes that cohere in complex compound words (Land 2013) and frequent consonant clusters, which make decoding a challenge. In isiXhosa, for example, these clusters can comprise up to five consonants in one syllable, for example in the word *i-ntshw-ela*. Consonants can be used together to form numerous different combinations, for example *hl, kh, mb, ny, ncw, ndl, tsw, tyhw*. Recent studies on benchmarking in the Nguni languages have shown that children struggle mostly with "complex consonant sequences" when reading (Ardington et al. 2020; Wills et al. 2022), justifying why these need to be taught explicitly and, more importantly, incrementally in Grade 1 and Grade 2. Numerous studies of Grade 1 children learning to read in languages with a transparent orthography such as German, Italian, Greek, Dutch, Portuguese, Turkish, and Finnish show mastery of reading skills after one year of instruction (Aro 2004). It can therefore be conjectured that orthography cannot be blamed for low literacy levels in African languages, but rather a lack of pedagogic linguistic understanding as to what constitutes correct reading instruction and reading materials for early literacy.

The research and theories above informed the conceptualisation and development of the Vula Bula readers.

2.3 Graded readers

Defining the term 'graded readers' can be difficult because it is wide-ranging and can be interpreted in different ways. Graded readers are usually described as being books of different genres containing simplified texts, typically graded according to vocabulary, complexity of grammar structures and number of words, to ensure accessibility to learners of the language. According to Mesmer (2001), decodable texts have two features: phonic regularity and lesson-to-text match. This is the systematic and explicit adherence

6 In conjunctive languages, "in writing, the affixes occur next to the … root … with mostly no orthographic spaces between the … affixes and the … root." (Schaefer, Probert & Rees 2020).

to a scientifically formulated scope and sequence of letters[7]. The text is then accompanied by a corresponding lesson, which explicitly teaches the graphemes before the children engage with the story (Dornbrack & Kazungu 2022). Basal readers are usually a levelled series of textbooks produced by an educational publisher. They focus on teaching reading either by a code-emphasis approach or a meaning-emphasis approach (Soler 2016). A code-emphasis approach relies heavily on phonemic awareness and decoding and word-attack skills. A basal-reading instruction programme uses texts that are written to teach reading, as opposed to using written texts to teach reading. This means that "the programme has been specifically designed to teach skills that have been proven to be helpful in learning to read, such as phonemic awareness, fluency, vocabulary, text comprehension, (including decoding and word-attack skills) and prosody" (Morin 2020).

Due to the transparent orthography and phonic regularity of African languages, Molteno uses a hybrid approach by combining elements of decodable texts and basal readers in the development of the Vula Bula series. These graded readers are therefore more similar to basal readers – readers that target specific language features and have a strong phonic focus. Since the mid-1990s, language teaching focus has shifted to skills acquisition, leading to a revival in the use of basal readers.

2.3.1 Graded readers: A critical resource for the Foundation Phase

The move to a whole-language approach to reading instruction from the late 1980s to the mid-1990s appears to have informed the generic definition of graded readers listed in the glossary of the CAPS (2011, 32) for English HL Grades R-3:

> *graded reading series: a series of readers divided in levels (e.g. Grade 1 - Level 1, 2 and 3). The levels go from easiest (simple vocabulary, short sentences, short text) to more difficult. Each level has a number of readers in it. The teacher assesses a learner and gets him/her started at the correct level. The learner reads several books at that level until he/she can read fluently at that level. The learner then moves up a level.*

However, when describing the methodology of group guided reading, which the DBE stipulates in the Foundation Phase, the CAPS describes texts that are usually contained in basal readers:

> *Graded readers will mostly be used for group reading. They should be at a lower level than the texts used for Shared Reading.* The text must be read with ease but still have a few challenges for the reader either at a decoding or at the comprehension level. The reader should recognise and quickly decode 90% - 95% of the words.
>
> (CAPS English HL Grades R-3, own emphasis)

The DBE recognises the importance of graded readers as a reading resource in that all its endorsed Foundation Phase literacy-research projects since 2015 have mandated the inclusion of graded readers (or graded reading anthologies) as a required component of the LTSM packs provided[8].

7 Numerous texts are analysed in order to determine the prevalence and frequency of phonic structures.

8 Projects include the Early Grade Reading Study 1 (EGRS I), Early Grade Reading Study 2 (EGRS II), Reading Support Project (RSP), and the Early Grade Reading Project (EGRP).

The background report for the 2030 Reading Panel (Spaull 2022) states that the only reading resource provided by the DBE to all schools are the DBE Workbooks – just one out of ten resources recommended for teaching reading in the Foundations for Learning Campaign: 2008-2011, gazetted by Minister G.N.M. Pandor in 2008 (see Figure 1).

Figure 1 **Government Gazette: Foundations for Learning Campaign: 2008-2011**

12 No. 30880 GOVERNMENT GAZETTE, 14 MARCH 2008

RECOMMENDED RESOURCES FOR LITERACY IN GRADES 1-3

For the walls	*For each learner*	*Resources for the teacher*
Alphabet frieze	A Workbook or Learners' book (to provide systematic development of literacy, phonic and language skills)	Big Books
Alphabet chart (letter, word, picture)	Set of small alphabet cards (for word building)	Graded readers
Birthday chart	Sight words on sheets/cards (to cut up and keep in a box – e.g. 25 words per term)	A list of reading words per reader
Weather chart	Un-lined Jotter	Read-aloud stories (for teacher)
Numbers word chart	Lined Handwriting book	Independent reading books (different topics, different levels)
Colour word chart	Lined Phonics book	A systematic programme which includes: • Phonemic awareness • Phonic development • Sight words • Language development
Vocabulary charts (words and pictures) e.g. • My Body Chart • My Family Chart • Fruit Chart • Farm animals • Actions	Lined Spelling book	(If one or other of these elements is missing, a supplementary programme will be needed). The programme(s) should include a workbook or learners' book for each learner – see below
Learners' names on flashcards	Lined Language/Writing book	
Sight words on flashcards	Personal Dictionary *(lined exercise book with letters of the alphabet on different pages)*	
	Wax crayons	
	Pencils	
	Sharpener	
	Soft eraser	
	Blank paper for drawing	

At the DBE's Teacher Development and Curriculum Management forum in 2019, the director of LTSM and Innovation at the DBE presented the strand: *Provisioning of LTSM* with the focus *Integrated Package*. This package included graded readers/anthologies as a mandatory/core/ 'must have' classroom reading resource.

Table 1 Recommended Resources for Literacy in Grades 1-3. DBE's Teacher Development & Curriculum Management Forum, 2019.

Recommended classroom pack of resources
Mandatory / Core / 'Must haves'
Foundation Phase
• Big Books (Grade 1 and 2) *(shared reading)*
• Graded BASAL readers/anthologies (Grade 1 and 2) *(Group Guided Reading; Independ)* ▪ readers in a carefully structured series, developmental, decodable, phonically sequential, varying levels of complexity
• Graded readers (Grade 3–7) ** All book packs to include both fiction and non-fiction, and different genres of texts*
• Puppets and masks *(can be teacher-made)*
• Dictionaries (bilingual) *(e.g. Lexicography Unit)*

In the *Right to Read and Write* report issued by the South African Human Rights Commission (2021), graded readers or reading anthologies are prominent again as a recommended resource for effective reading instruction:

> *All Grade 1 and 2 learners should receive two anthologies of graded-reader texts per year (20 titles per anthology so 40 titles for the year, one per week) for Group Guided Reading, Paired Reading, and Independent Reading. These readers should be a carefully structured series: developmental, decodable, phonically sequential, and of varying levels of complexity. All Grade 1, 2 and 3 learners should receive graded-reader texts in their home language.*

In addition, this report advocates the use of Vula Bula as an economical solution:

> *The possibility of utilising Open Access materials (for example the Vula Bula series) should also be explored since this is a viable way of getting high-quality materials to all learners in their Home Language and within existing budget constraints.* (2021, 30)

3 Molteno's Vula Bula: Development

To begin the development process, Molteno compiled a core vocabulary list suitable for Foundation Phase learners in isiXhosa. The texts for the Vula Bula Grade 1 and Grade 2 readers were drawn mainly from this list. Words were carefully selected strictly on a phonics basis, with new consonant clusters (blends, digraphs, trigraphs, etc.) introduced incrementally, and systematically. The readers were divided into levels of difficulty from one star to six stars, with each level building on the phonic reading practice of the previous level(s), and supporting different stages of reading development from early emergent reading to fluent reading. All the stories at a particular star level need to be mastered – read fluently with understanding – before the learner can progress to the next star level.

Because African languages in South Africa have transparent orthographies, all words *can* be sounded out, and therefore do not need to be learned by sight. The concept of sight words is different from that in English, which has an opaque orthography. In English, words such as "one", "here", "are", "love", etc. *cannot* be sounded out and must be learned as whole words, that is, by sight. Many of these words are common and are therefore counted as high-frequency words. That said, there are also high-frequency words in African languages that need to be recognised early on in order to build fluency, and which are also included in story texts in order to maintain reader interest.

Using a corpus-based approach, Molteno has developed authentic, empirically proven high-frequency word lists in all nine African languages. Word tokens were taken from texts written specifically for children. It is notable that a number of common words in African languages contain a complex consonant cluster, necessitating exposure to these words as 'sight words' in very early readers, for example *incwadi* (a book) in isiXhosa.

Figure 2 Vula Bula 3-star isiXhosa reader: Sidlala undize

Igama elitsha amabaliqhele (ukuqwalasela kwakhona)
lixesha

Amagama amatsha amabawaqhele
yho | ephi

Izandi (ukuqwalasela kwakhona)
hl, kh, nd, ng, th

Izandi
dl (sid**l**ala, uku**dl**ala, ba**dl**ale, si**dl**ale, ngum**dl**alo, lesi**dl**o)
nk (u**Nk**ululeko, no**nk**e, na**nk**o, so**nk**e)

VulaBula Uthotho lweencwadana zebanga loMgangatho osisiSeko
Ipapashwe 2013 ngabakwa-Molteno Institute for Language and Literacy
Zisekwe kwimvelo yothotho lweencwadana exhaswe yi-Zenex Foundation

★★★ Sidlala undize
- Umhleli wothotho: Jenny Katz
- Abaqambi bamabali: Jenny Katz noMirna Lawrence
- Umsebenzi: Jenny Katz
- Incutshe kulwimi isiXhosa: Nolitha Bikitsha
- Umzekelisi: Sandy Lightley
- Uyilo nongqaleko: Resolution

© 2013 Molteno Institute for Language and Literacy

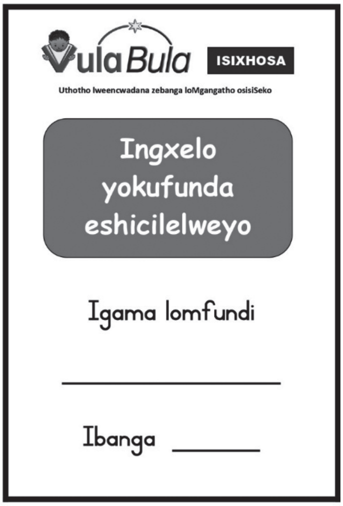
ISBN 978-1-77580-141-2

Creative Commons Attribution-NonCommercial-NoDerivatives (CC BY-NC-ND) 4.0 International Licence

Disclaimer: You are free to download and share this work as long as you attribute the Molteno Institute for Language and Literacy, but you may not change this work in any way or use it commercially.

On the inside front cover of each reader, sight words taken from that text are listed (Figure 2). These may be high-frequency words, or they may be hard-to-read common words that contain unfamiliar phonic structures. Revision sight words that appear in the text that have been introduced in previous star levels are also included. Finally, all words containing the specific consonant cluster focus in the text are listed. Consonant clusters that appear in the text but have already been introduced in previous star levels, are also revised.

Figure 3.1 and 3.2 **Vula Bula Grade 1, Grade 2 isiXhosa Reading Record**



IINCWADI ZESIXHOSA	AMAGAMA AMATSHA AMABAWAQHELE	IZANDI	UMHLA NOMQONDISO NGOKUTYIKITYA NGUTITSHALA	UMHLA NOMQONDISO NGOKUTYIKITYA NGUMZALI
INQANABA loku-✶				
Hayi!	uyacinga	ukuziqhelanisa izandi zonoobumba be-alfabethi		
Vula vala	iphi			
Baleka!	ndoda			
Lala	lixesha			
Bala	futhi			
Coca				
Zoba usike	mnye			
Mamela		nd		
INQANABA lesi-✶✶				
Jika	kakhulu	ukuziqhelanisa izandi zonoobumba be-alfabethi		
Utata usele nosana	thula			
Umnikelo	enkosi			
Lelikabani eli vili?	kakuhle, enkosi			
Uhambo	kakhulu	UKUQWALASELA KWAKHONA: nd NTSHA: mb, ng		
Sebeza	benje, namhlanje	UKUQWALASELA KWAKHONA: nd NTSHA: hl, th		
Ekhaya	apha, bonke, lilizwe	hl, kh		
INQANABA lesi-✶✶✶				
Ncedani!	ukhwela	UKUQWALASELA KWAKHONA: th NTSHA: nc		
Yophukile ifestile	ifestile	UKUQWALASELA KWAKHONA: nd NTSHA: bh, ph		
Uphi uZinzi?	uphi	UKUQWALASELA KWAKHONA: kh, nd, ng NTSHA: nz, sh		
Sityiwe isonka	efestileni, kodwa, uphinde	UKUQWALASELA KWAKHONA: kh. nd NTSHA: nk, ty		
Sidlala undize	UKUQWALASELA KWAKHONA: lixesha NTSHA: yho, ephi	UKUQWALASELA KWAKHONA: hl, kh, nd. ng. th NTSHA: dl, nk		
Izinyo	encwadini	UKUQWALASELA KWAKHONA: ng NTSHA: lw, ny		

Source: https://vulabula.molteno.co.za/sites/default/files/readers/OER%202015%20Reading%20Re-cord%20ISIXHOSA.pdf

A Reading Record for each learner accompanies the Grade 1 and Grade 2 series of readers (Figure 3.1). The Reading Record is used to monitor the reading progress of each individual learner. The teacher uses the Reading Record to document whether a learner has managed to master all the stories at a particular star level.

The Grade 3 readers contain more pages, more text per page, longer sentences, greater variation in sentence pattern, more formal and descriptive language, and richer vocabulary (Figure 3.2). Reading now becomes more automatic as learners approach reading independence, with the focus shifting from decoding skills to fluency and comprehension. The readers are also levelled in terms of difficulty, taking into account text length, amount of information provided, and conceptual difficulty.

Each story is preceded by a list of vocabulary words and a list of hard-to-read words, which the learners must understand and practise decoding before they read the story.

Each story is followed by a set of comprehension questions, a language activity, and a writing activity (Figure 4). The comprehension questions are modelled on questions used in the DBE's Annual National Assessments (ANAs) and in the PIRLS assessments, including closed and open-ended questions, multiple choice, sequencing events in the story, and true and false. The language and writing activities are specifically linked to the Grade 3 Home Language CAPS and the DBE Grade 3 Workbooks for each African language. Finally, a model-answers book is provided for the teacher, containing the answers to the comprehension questions and language activities in the readers, and providing exemplars for the writing activities.

Figure 4 **Vula Bula Grade 3 isiZulu Reader 3B: after reading comprehension questions, language activity and writing activity**

Awu, Mnumzane Nkawu!

A. Ukuqondisisa

Phendula imibuzo ngemisho ephelele.

1. Walungisani ezohambisana netiye uNkosazana Nkawu?
2. UNkosazana Nkawu wabeka isinkwa eduze kwefasitela elivulekile __________.
   a) ukuze UMnumzane Nkawu afikele kuso
   b) ukuze UMnumzane Nkawu asihogele
   c) ukuze siphole
3. Wenzani uMnumzane Nkawu ukuze uNkosazana Nkawu amthande? (*Bheka ekhasini 23.*)
4. Kungani uNkosazana Nkawu athatha isikhathi eside ukuzilungisa?
5. UMnumzane Nkawu wavele wadla isinkwa noshizi ngoba __________.
   a) akakwazanga ukuzibamba ebona isinkwa esinuka kamnandi
   b) wayesekhathele ukulinda uNkosazana Nkawu
   c) wayeselambe kakhulu
   d) Konke lokhu okungenhla.
6. Wenza kanjani uNkosazana Nkawu ukuthi abukeke?
7. Ucabanga ukuthi kwakufanele uNkosazana Nkawu amthukuthelele kangaka uMnumzane Nkawu? Nikeza isizathu sempendulo yakho.

29

8. Beka isiphambano (X) eduze kwemisho eyiphutha, uthikhe (✓) eduze kwemisho eliqiniso.

a)	Izinkawu zinemisila emide.	
b)	Izinkawu ziyakwazi ukubhaka.	
c)	Izinkawu zinezandla ezifana nezethu.	
d)	Izinkawu zinezingubo zokugqoka.	

B. Ukusetshenziswa kolimi

DBE Ibanga lesi-3 Incwadi yoku-1: ikh. 92
DBE Ibanga lesi-3 Incwadi yesi-2: amakh. 41, 46, 75, 112

Khetha <u>futhi</u> noma <u>kodwa</u> ukuhlanganisa le misho ehamba ngamibili.

1. UNkosazana Nkawu usika ushizi. Uwubeka epuletini.
2. UMnumzane Nkawu ufuna ukudla. UNkosazana Nkawu akaqedi.
3. UNkosazana Nkawu ufaka ubucwebecwebe. Ulungisa izinwele.
4. UNkosazana Nkawu uthukuthele kakhulu. UMnumzane Nkawu usuthi ujabulile.

C. Masibhale

DBE Ibanga lesi-3 Incwadi yoku-1: amakh. 36, 38
DBE Ibanga lesi-3 Incwadi yesi-2: amakh. 18, 86, 106, 109

Yenza sengathi ungu Mnumzane Nkawu. Bhala <u>incwadi</u> iye kuNkosazana Nkawu uxolise kuye ngokwenzekile.

30

Source: Vula Bula Grade 3 isiZulu Reader 3B, pages 29 and 30. From: https://vulabula.molteno.co.za/reader/grade-3-incwadi-3b-isizulu

4 Molteno's Vula Bula: Implementation

In 2010, the Gauteng Department of Education (GDE) embarked on the Gauteng Primary Literacy Strategy (GPLS)[9] in over 800 schools. In 2012, the South African Institute for Distance Education (Saide) compiled a GDE-authorised evaluation report in which the quality, suitability, and efficacy of the graded readers originally selected for the project were reviewed and found lacking. The GDE then commissioned Molteno to modify its existing Grade 1 isiXhosa graded readers into a stand-alone product for group guided reading, and to develop a similar series in seven additional African languages (isiNdebele, isiZulu, Sepedi, Sesotho, Setswana, Tshivenḓa, and Xitsonga). This set of new graded reading series was successfully completed in 2013.

Having taken 18 months to complete the series of 32 isiXhosa readers, Molteno recognised that there was insufficient time to develop and illustrate original stories for each new language in the time period required by the GDE. Due to the similar orthographies of all the African languages, Molteno decided to try to version the readers instead. First, the stories were translated into each new African language, and words containing complex consonant clusters were removed by paraphrasing or modifying the story. A phonic focus was then identified, and finally, as many additional words as possible containing the target consonant cluster were added to the text without making the story contrived. This versioning process required many specialist skills; it was difficult, laborious, and time-consuming, but ultimately successful.

After receiving reports from GPLMS schools in 2014 that the readers were 'too easy' for Grade 3 learners, the GDE then commissioned Molteno to adapt all 32 titles for suitable reading practice in Grade 3. The Vula Bula series was thereby extended to suit the whole Phase.

Since the GPLMS in 2013, the Vula Bula graded readers have been successfully provided to many large-scale and important Foundation Phase projects such as Jika iMfundo in KwaZulu-Natal (in isiZulu), the EGRS I, RSP and EGRP in the North West (in Setswana), and in national NECT interventions.

From 2016 to 2019, with a grant from the Zenex Foundation, Molteno also developed core Foundation Phase LTSM in all nine African languages to support reading, including alphabet and phonics classroom wall-friezes and cards, high-frequency wordlists, and Big Books. In 2021, Molteno also completed the development of the graded-reader series in Siswati.

Part of the grant agreement involved making Vula Bula available as an open-education resource (OER). Vula Bula is published by Molteno using a Creative Commons-Attribution-Non-Commercial-No Derivatives licence (CC BY-NC-ND). This license is the most restrictive of the six main OER licenses: allowing the work to be downloaded and shared with others as long as the creators are credited, but the work may not be changed in any way or used commercially. This is necessary in order to protect the structure and sequence of the graded texts in each series, and to acknowledge Molteno as the intellectual-property owner of all the Vula Bula resources. The Vula Bula website is zero-rated, allowing free online access and downloading of resources.

9 The GPLS later included Mathematics in its scope and was renamed the Gauteng Primary Literacy and Mathematics Strategy (GPLMS).

In 2019, at the request of the Eastern Cape Department of Education, Molteno re-packaged its sets of Vula Bula isiXhosa and Sesotho graded readers into graded-reading anthologies for Grades 1, 2, and 3. These were supplied to all the Foundation Phase learners in the province. Anthologies are more cost-effective to print than single readers, and allow for easier classroom management. The impact of this provision has delivered positive results in increasing isiXhosa learners' reading rate of words per minute (Ardington & Spaull 2022). Molteno has since completed the consolidation of all its separate readers into graded-reading anthologies for Grades 1, 2, and 3 in all nine African languages.

5 Reader text analyses and findings

5.1 Analysing isiXhosa texts

In an effort to ensure that the Vula Bula texts were at an appropriate instructional level, and to determine a viable blueprint for the development of early readers that could effectively help to improve current deficient literacy levels in African languages, an intensive analysis of two published isiXhosa graded-reading schemes, namely Izinga Eliphezulu and Siyakhula was done in 2014 (Katz 2019).

The texts were analysed in terms of the total number of sentences and words per text; number and type of words per sentence; number of syllables and graphemes per word; number of graphemes and their relative complexity per word; number and type of words per sentence.

It was found that the isiXhosa texts were translations of English versions and therefore did not function to systematically teach decoding skills at the appropriate instructional level (Katz 2019). Significantly, the analysis proved, across all 60 texts, that the average isiXhosa word contained approximately three syllables. This means that, even in early readers, learners will need to be able to decode mostly long words made up of at least three syllables. We realised that for a conjunctive language like isiXhosa it would be difficult to avoid polysyllabic words of up to six syllables; the only word structure that could be carefully controlled during development was the exclusion of complex consonant sequences containing two to five consonants. In Table 2, the average shorter word length (5.9 letters) for the Vula Bula readers can be explained by the absence of these consonant clusters.

Table 2 Average of number of letters and syllables per word across three isiXhosa graded reading series

ISIXHOSA GRADE 1 SERIES	Total average number of syllables per word	Total average number of letters per word
Izinga Eliphezulu	3.3	7.1
Siyakhula	3.2	7.0
Vula Bula	3.0	5.9

More recently, Molteno undertook a similar analysis of the texts in four Grade 1 isiXhosa DBE graded-reader texts (Readers 3, 4, 5 and 6)[10], and compared these with the same readers from the Grade 1 series in the previous analyses (Izinga Eliphezulu [IE], Siyakhula [S] and Vula Bula [VB]. The complexity of the phonic structures therein is summarised in Table 3:

Table 3 Average of number of complex consonant clusters across four isiXhosa graded reading series

Averages across DBE Grade 1 Graded Readers 3-6, Izinga Eliphezulu (IE) Grade 1 Readers 3-6, Siyakhula (S) Grade 1 Readers 3-6 and Vula Bula (VB) Grade 1 Readers 3-6				
Language XHOSA	**Average no. of words per text**	**Average no. of 2-letter consonant blends & digraphs per text**	**Average no. of 3-5 letter consonant clusters per text**	**Phonics structures (consonant clusters) included in the four texts**
DBE XHOSA Readers 3-6	77.5	11.25	3.25	lw, nd, hl, th, nj, ny, nk, ng, ty, kw, nw, dl, bh, ph, ts, kh, mb, mv, yh, nq, mp, gq, dw, nx, nc, ch, qh, tw tsh, ndl, ngw, tyw, nts, ntw, thw, khw, ngq
IE XHOSA Readers 3-6	43.75	12.25	2	bh, dl, hh, hl, kh, kw, lw, mh, mv, nc, nd, ng, nj, nk, nt, nx, ny, nz, ph, qh, sh, th, xh, zw khw, ngc, ntl, nts, tsh, tyh
S XHOSA Readers 3-6	28	7.75	1.75	bh, hl, jw, kh, kw, lw, mb, mv, nd, ng, nj, nk, nq, nw, ny, ph, sh, sw, th ndl, nkw, nts, qhw, tsh
VB XHOSA Readers 3-6	33	0.75	0	nd, sh, th

It is evident from the averages and lists shown that the texts in the first three sets of readers demand advanced phonic knowledge and decoding ability from Grade 1 learners in terms of complex consonant clusters, whereas the Vula Bula readers require learners to recognise just one common consonant blend (*nd*) and two common consonant digraphs (*sh*, *th*). This clearly shows the difference between texts that are written to teach reading (Vula Bula), as opposed to written texts translated from English to teach reading (the other three series).

5.2 Analysing additional African language texts

An additional analysis was done on the same four DBE Grade 1 readers but this time extending to three more African languages: Sepedi (Sotho language family), Tshivenḓ a, and Xitsonga. The complexity of the phonic structures in these Grade 1 Readers 3-6 was then compared with the Vula Bula Grade 1 Readers 3-6 in the same languages, as shown in Table 4:

10 The story text analysis began with DBE Level 1 Book 3 – Level 1 Book 1 contained no text except for the title, and Level 1 Book 2 was not a coherent text, but rather a selection of illustrated sentences.

Table 4 Average of number of complex consonant clusters between DBE and VB readers across four languages

Averages across DBE Grade 1 Graded Readers 3-6 and Vula Bula (VB) Grade 1 Readers 3-6				
	Average no. of words per text	**Average no. of 2-letter consonant blends & digraphs per text**	**Average no. of 3-5 letter consonant clusters per text**	**Phonics structures (consonant clusters) included in the four texts**
DBE XHOSA Readers 3-6	77.5	11.25	3.25	lw, nd, hl, th, nj, ny, nk, ng, ty, kw, nw, dl, bh, ph, ts, kh, mb, mv, yh, nq, mp, gq, dw, nx, nc, ch, qh, tw tsh, ndl, ngw, tyw, nts, ntw, thw, khw, ngq
VB XHOSA Readers 3-6	33	0.75	0	nd, sh, th
DBE SEPEDI Readers 3-6	125.25	9.25	3	ng, tš, ny, ph, kg, tl, sw, ts, th, mm, hl, nn, gw, rw, kh, kw, lw, mp, bj tšh, nny, kgw, ngw, tšw, tsh, mpš, psh, ntl ntsh, ntlw
VB SEPEDI Readers 3-6	45.75	0.75	0	tš, ts, tl
DBE VENḒA Readers 3-6	121	13.25	3.5	ṱh, nd, dz, sh, zw, ph, ng, mb, vh, kh, nṋ, fh, nz, nḓ, rw, nt, th, nn, pf, lw, sw, ts, mm tsw, nṱh, dzh, tsh, mmb, nzh, nnḓ, nnd, dzw mmbw
VB VENḒA Readers 3-6	63.75	3.5	0.5	vh, nd, ng, fh, kh, dz, ph, sh tsh
DBE TSONGA Readers 3-6	118.5	9.5	5.5	rh, lw, sw, dy, nh, ph, n'w, tl, ng, ts, hl, mb, pf, ny, nt, th ndz, nhl, ndl, nkw, tsh, mby, gqh, ngh, tlh ntlh, tshw ntshw
VB TSONGA Readers 3-6	55.25	2.25	0.25	pf, n'w, ng, hl, tl, by ndl

The analyses revealed that the DBE Grade 1 readers could not be described as a developmental, structured series in terms of decodability and incremental phonic sequencing. It is evident that these are translations of the English home-language graded readers. Grade 1 learners are expected to read words containing many complex consonant clusters at a time when they are only just beginning to grasp the concept of phoneme-to-grapheme-matching. Simply, Grade 1 learners will not be able to recognise and quickly decode 90% to 95% of the words.

Translation from English into African languages in the early grades is usually undertaken without awareness that short, common, easily decodable words in English often transmute into considerably longer, multisyllabic words containing several affixes and complex consonant combinations, making a text considerably more challenging to read. The (in)effectiveness of English-to-African-language translation in producing early grade readers interrogated in this chapter has shown that the way in which early-reading materials for young learners in African languages have been developed to date is contributing significantly to the literacy crisis. Approaches are being used for the teaching of reading in African languages that may work successfully in English, but which are not well suited to agglutinative languages, and especially not those with conjunctive orthographies.

There is a critical need for publishers to produce more storybooks in African languages to enable children to harness their existing knowledge of life and oral language in order to make meaning, rather than simply decoding the text (Edwards & Ngwaru 2012). Good translations of good stories will serve this purpose, and these do exist in the DBE's current National Catalogue. However, if being able to decode text fluently using reading strategies such as phonic synthesis and syllabification is a critical literacy skill in Grade 1 and Grade 2 – especially in agglutinative, conjunctive African languages – we propose that these should form the theoretical basis for developing early graded reader series. As evidenced by the Vula Bula readers, phonic structures are a text feature that can be manipulated and controlled to allow the mechanics of reading to be introduced in an incremental, systematic way to ensure sustainable and lasting reading competence.

6 Impact of the Vula Bula readers

All nine Vula Bula graded reader series have not been vetted for inclusion in the National Catalogue as they were developed after submissions closed in 2011 (have not been reopened to date). The readers have however undergone both qualitative and quantitative evaluations, some of which are summarised below:

> *Teachers reported that the books were very beneficial to the learners and that they have made the teaching of reading very manageable and enjoyable ... The books inject a love of reading in the learners. Learners who struggle to read are given readers to match their ability, and then are gradually moved up as they master each level. In this way learners develop confidence in their own reading; in turn this increase in confidence then leads to success in reading and generates the enthusiasm to read more ... Each level provides learners with a wealth of vocabulary required for the acquisition of reading skills as well as for reading comprehension which is critical for successful learning. The books have contributed greatly to improved reading in the Foundation Phase (Grades 1-3). Due to the intensity of the Vula Bula reading programme, learners are able read fluently in their own mother tongues.*
>
> Ms Kholofelo Makhaga, Directorate: Curriculum Coordination – GET, Office of the GPLMS, 2015

In 2013, Saide conducted a desktop evaluation of Molteno's Setswana Vula Bula Grade 1 and Grade 2 graded readers. The evaluation concluded as follows:

> *The 32 Vula Bula Setswana readers are an excellent example of a graded-reading series. Unlike most of the material produced in African languages, the series:*
> - *Supports the balanced approach in the teaching of reading;*
> - *Systematically introduces and practises the phonics appropriate for each level in the language of the readers;*
> - *Shows how the simplest readers can still be stories and engage the learners;*
> - *Demonstrates how non-fiction topics can be written in a way that encourages reading for pleasure;*
> - *Contains illustrations that can be used by the imaginative teacher to encourage critical thinking;*
> - *Has content that is generative – a range of activities can and have been developed around the books.*

In 2014, Molteno collaborated with Mindset Network on their Mindset Teach Ukusiza project, part of the School Capacity and Innovation Project (SCIP) funded by USAID. Project schools were situated in the Free State, KwaZulu-Natal, North West, and Mpumalanga. Molteno supplied Vula Bula readers in the relevant African languages to half of the project schools. An evaluation by Africa Strategic Research Corporation in 2015 suggested gains in four EGRA literacy sub-tasks for learners who received the Vula Bula readers compared to those who did not.

7 Conclusion and way forward

Research shows that becoming a fluent reader in one's home language can accelerate the acquisition of a more complicated additional language (Abadzi 2006). Without a sound grounding in their home language, children struggle to achieve language success in either their first or second language by the end of the Foundation Phase. This resonates in the South African context where English becomes the language of learning and teaching in Grade 4.

The Vula Bula graded readers were developed using a phonics-based approach, combining structural features of both decodable texts and basal readers to best suit the transparent orthographies of African languages. Great effort was invested in creating enjoyable stories that would keep children interested in turning the pages while practising their reading skills. This is an arduous and complicated development process that demands a multitude of specialised skills – perhaps the reason why there are no other similar reading series currently available – but it is well worth the effort for its valuable contribution towards achieving early grade reading success in African languages.

With regard to the agglutinative nature of African languages, particularly the conjunctive languages, the authors recognise that syllabification is an important skill for developing reading fluency and comprehension, and we believe that further research is needed to measure whether awareness of the morphological structure of

words can also improve reading fluency and comprehension.
The Covid-19 pandemic has exacerbated South Africa's existing early grade literacy woes, and authorities need to embrace bold decision-making and action to improve the situation. In 2019, President Cyril Ramaphosa stated that "every ten-year-old child will be able to read for meaning within the next five years" (SONA 2019). An urgent national commitment to provide reading resources such as Vula Bula to all Foundation Phase learners can help to achieve this.

References

Abadzi, H. (2006). Efficient Learning for the Poor. *Insights from the Frontier of Cognitive Neuroscience*. Washington DC: World Bank Publications, 36–57.

Ardington, C., Wills, G., Pretorius, E., Deghaye, N., Mohohlwane, N., Menendez, A., Mtsatse, N. and Van der Berg, S. (2020). *Summary Report: Benchmarking early grade reading skills in Nguni languages*. Stellenbosch: ReSEP, Stellenbosch University. Cape Town: SALDRU, University of Cape Town. Chicago: NORC at the University of Chicago.Pretoria: Department of Basic Education.

Ardington, C. & Spaull, N. (2022). Sequenced for success: Evaluating the impact of openly licensed Vula Bula anthologies of graded readers on early grade reading outcomes in South Africa. In Spaull, N. & Taylor, S. (Eds) *Early Grade Interventions in South Africa*. Cape Town: Oxford University Press.

Aro, M. (2004). Learning to read: The effect of orthography. *Jyväskylän yliopisto*. University of Jyväskylä. 4(36), 39–44.

Aro, M. & Wimmer, H. (2003). Learning to read: English in comparison to six more regular orthographies. *Applied Psycholinguistics*. 24(04), 621–635.

Clay, M.M. (1991). *Becoming literate: The construction of inner control*. Portsmouth: Heinemann.

Cunningham, A. & Stanovich, K. (2003). Reading can make you smarter! *Principal*, 83, 34–39.

Department of Basic Education (2008). *National Reading Strategy*. Pretoria: DBE.

Department of Basic Education. (2011). *Curriculum and Assessment Policy Statement (CAPS)*. Pretoria: DBE.

Diemer, M., van der Merwe, K., & de Vos, M. (2015). The development of phonological awareness literacy measures for isiXhosa. *Southern African Linguistics and Applied Language Studies*. 33(3), 325–341.

Dixon, M. & Stuart, M. & Masterson, J. (2002). The relationship between phonological awareness and the development of orthographic representations. *Reading and Writing*. 15, 295–316.

Donaldson, M. (1978). *Children's Minds*. Glasgow: Fontana/Collins.

Dornbrack, J. & Kazungu, T. (2022). Room to Read's use of decodable stories in their structured literacy programme In Spaull, N. & Pretorius, E. (Eds) *Early Grade Reading in South Africa*. Cape Town: Oxford University Press.

Edwards, V. & Ngwaru, J.M. (2012). African language books for children: issues for authors. *Language, Culture and Curriculum*, 25(2), 123–137.

Fountas, I.C. & Pinnell, G.S. (1996). *Guided reading: Good first teaching for all children*. Portsmouth, NH. Heinemann.

Heath, S.B. (1982). What No Bedtime Story Means: Narrative Skills at Home and School. *Language in Society*, 11(1), 49–76.

Howie, S.J., Combrinck, C., Roux, K., Tshele, M., Mokoena, G.M., & McLeod Palane, N. (2017). *PIRLS Literacy 2016 Progress in International Reading Literacy Study 2016: South African Children's Reading Literacy Achievement*. Pretoria: Centre for Evaluation and Assessment.

Izinga Eliphezulu isiXhosa Grade 1 reading series (2006). Johannesburg: READ Educational Trust.

Janks, H. (2011). Making sense of the PIRLS 2006 results for South Africa. *Reading and Writing*. 2(2), 27–39.

Katz, J.L. (2019). Investigating the Linguistic Effectiveness of Early Reading Schemes in isiXhosa: A Phonological and Orthographical Analysis of Three isiXhosa Grade 1 Graded Reader Series. Rhodes University. Masters Thesis. Available at: http://hdl.handle.net/10962/92798 (Accessed 14 November 2022).

Kgobe, M., Baatjes, I., Sotuku, N. (2012). *Report of The Community Literacy and Numeracy Group Project (CLING)*.

Land, S. (2013). Tracking reading in isiZulu. Unpublished draft of doctoral thesis. Centre for Adult Education. University of Kwazulu-Natal. Cites from Land, S. (2011). Open to the flaw: Comparing the different skills required for reading English and isiZulu. *Language Matters*, 42(1), 50–68.

McDougall, S., Brunswick, N., & de Mornay Davies, P. (2010). Reading and dyslexia in Different orthographies: An introduction and overview. In McDougall, S., Brunswick, N., & de Mornay Davies, P. (Eds), *Reading and Dyslexia in Different Orthographies*, Hove: Psychology Press.

Mesmer, H. (2001). Examining the theoretical claims about decodable texts: Does text decodability lead to greater application of letter-sound knowledge in first grade readers? In *Yearbook: National Reading Conference* 50, 444–459.

Morin, A. (2020). How the Basal Reading Instruction Program Works. *Verywellfamily.com*. Available at: https://www.verywellfamily.com/reading-programs-what-is-basal-reading-instruction-620997 (Accessed 21 February 2022).

National Education, Evaluation and Development Unit (NEEDU). (2013). *National Report 2012: Summary Report*. Pretoria.

National Institute of Child Health and Human Development (2000). Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction. *Report of the National Reading Panel*. (NIH Publication No. 00-4769). Washington, DC: U.S. Government Printing Office.

Oakhill, J. (1993). Developing skilled reading. In Beard, R. (Ed) *Teaching literacy balancing perspectives*. London: Hodder & Stoughton.

Pearson, P.D. (2004). The Reading Wars. *Educational Policy*, 18(1), 216–252. Reprinted by permission of Sage Publications, Inc.

Pelli, DG. & Tillman, K.A. (2007). Parts, Wholes, and Context in Reading: A Triple Dissociation. *PLoS ONE*. 2(8).

Ramaphosa, C. (2019). Vision 2030. *State of the Nation Address (SONA)*. 20 June. Cape Town.

Saide (2012). Materials evaluation of Foundation Phase Literacy Resource Packages Gauteng Primary Literacy Strategy. Unpublished report commissioned by the Gauteng Education Development Trust.

Saide (2013). Evaluation of the Molteno Setswana Grade 1 Vula Bula Graded Readers. Unpublished report commissioned by the Molteno Institute for Language and Literacies.

Schaefer, M., Probert, T., & Rees, S. (2020). The roles of phonological awareness, rapid automatised naming and morphological awareness in isiXhosa. *Per Linguam*. 36(1), 90–111.

Siyakhula isiXhosa Grade 1 reading series (2009). Johannesburg: Oxford University Press.

Snow, C., Griffin, EP., & Burns, MS. (2005). *Knowledge to Support the Teaching of Reading*. San Francisco, CA: John Wiley & Sons/Jossey-Bass.

Soler, J. (2016). The politics of the teaching of reading. *Prospects*. (46), 423–433.

South African Human Rights Commission (2021). *The Right to Read and Write*. Cape Town.

Spaull, N. (2022). *2022 Background Report for the 2030 Reading Panel*. Cape Town.

Vula Bula graded reading series (2022). Johannesburg: Molteno Institute for Language and Literacies.

Wills, G., Ardington, C., & Sebaeng, L. (2022). Foundational skills in home language reading in South Africa: Empirical evidence from 2015–2021. In Spaull, N. & Pretorius, E. (Eds) *Early Grade Reading in South Africa*. Cape Town: Oxford University Press.

Wolf, M. (2008). *Proust and the Squid: The Story and Science of the Reading Brain*. Cambridge: Icon Books Ltd.

Ziegler, J. & Goswami, U. (2006). Becoming literate in different languages: similar problems, different solutions. *Developmental Science* 9(5). 429–453.

07

Sequenced for success: Evaluating the impact of openly licensed Vula Bula anthologies of graded readers on early grade reading outcomes in South Africa

CALLY ARDINGTON & NIC SPAULL

Abstract

In this chapter we evaluate the large-scale distribution, uptake and impact of a new reader: The open access Vula Bula graded-reading anthologies comprise a collection of 44 phonically sequenced levelled stories for Grades 1 and 2 and a collection of 22 grade-level stories for Grade 3. In 2019 and 2020, the Eastern Cape Department of Education distributed the anthologies to all Grade 1-3 learners in the province at a cost of approximately R15 ($1) per anthology. Drawing on a combination of teacher, principal, and learner interviews and classroom audits across 96 schools, we provide a detailed overview of the distribution, uptake, and use of these anthologies. Our evidence points to successful distribution and high usage of the anthologies both in the classroom and at home. We exploit variation in exposure to the anthologies within-grade-between-schools, and within-schools-between-grades to provide quasi-experimental evidence of the causal impact on reading outcomes in the home language (isiXhosa). We find that cohorts exposed to the anthologies read three additional isiXhosa words correctly per minute and that this impact represents a 20% improvement over comparable cohorts over the same period.

KEYWORDS
graded readers,
open access materials,
EGRA,
Vula Bula,
Eastern Cape

1 Introduction

Locally and internationally, there is now widespread agreement among researchers and policymakers that all children ought to be in school, and they should learn how to read and write for meaning and pleasure (Moore et al. 2017). In South Africa, both the president and the minister of Basic Education have stated that "reading for meaning by the age of 10" is now an apex priority for the country, and these statements have been codified in the key policy documents of the Presidency (DPME, 2020) and the Department of Basic Education (DBE, 2020). Yet in 2016 only 22% of Grade 4 children could reach the PIRLS Low International Benchmark (Howie et al, 2017). While this has been improving over time from 13% in 2006, current projections suggest that by 2031 only 36% of South African Grade 4 learners will be able to read for meaning (Spaull 2022).

There are numerous reasons for this low level of achievement, and many of these are multifaceted and complex, relating to pedagogy, resourcing, training, and accountability. However, some of the reasons are more straightforward: most South African classrooms lack appropriate texts to teach children to read. The current study reports on a province-wide initiative to provide cost-effective texts to children, helping them learn to read.

In this chapter we evaluate the large-scale distribution, uptake, and impact of a new format of reader: the open access Vula Bula graded reading anthologies, published by the Molteno Institute for Language and Literacies. This is essentially a series of three anthologies – one for each grade – made up of a collection of phonically-sequenced levelled stories from Story 1 at the start of Grade 1, to Story 44 at the end of Grade 2 and a collection of 22 grade-level stories for Grade 3. Anthologies were distributed to all Grade 1-3 children in one South African province (the Eastern Cape) in 2019 and 2020. The programme was initiated and funded by the Eastern Cape Department of Education (ECDOE). For 2019, 825,000 readers were delivered to 463,276 children across 4,298 schools at a cost of approximately R15 ($1) per anthology. We use data from 168 teacher interviews, 876 learner interviews, 93 principal interviews, and 77 classroom audits to provide a detailed overview of the distribution, uptake, and use of these anthologies. We then provide quasi-experimental evidence of the impact of this programme on reading outcomes using repeated Early Grade Reading Assessment (EGRA)-type data from 1984 learners in 85 no-fee schools. Using a difference-in-difference (DD) and difference-in-difference-in-difference (DDD) methodology we exploit variation in exposure to the anthologies within-grade-between-schools, and within-schools-between-grades to identify the causal impact on reading outcomes in the learners' home language (isiXhosa).

We believe the findings presented here provide strong evidence that this new resource for teaching reading (using open-access home-language anthologies of graded readers at a ratio of one book per child) is a cost-effective innovation that will help South Africa, and the rest of the continent, move more quickly towards the universal acquisition of basic reading skills.

2 Background and literature

In many developing countries access to learning and teaching support materials is often woefully inadequate, and this is especially so in the case of reading books and other text materials (UNESCO, 2016). While children can learn letters and words through communal instruction with a teacher writing on a board, learning to read is an individual activity that cannot be developed in the absence of books. Given this, it is perhaps unsurprising that in a review of 18 early grade reading programmes in developing countries, Graham and Kelly (2018) identify "providing supplementary instructional materials" as one of five common components across the interventions. Similarly, Stern et al. (2021, 3) report that one common feature across large-scale successful early grade reading programmes in developing countries was that "student books were available at a 1:1 ratio for all students". While materials provision by itself is not a sufficient condition for learning to read, the literature suggests that it is increasingly thought to be a necessary condition (Evans & Acosta 2021).

Before reviewing the quantitative literature on the impact of books on learning outcomes in developing countries, it is worth emphasising that the type and quality of books is rarely mentioned in evaluations of those books[1]. Notwithstanding this, it is still useful to review the evidence of textbooks on learning outcomes generally, and the impact of books (story books, readers, workbooks etc.) on early grade reading outcomes specifically.

2.1 International evidence of the impact of textbooks or books

Several studies in developing countries published in the 1980s and 1990s have shown positive associations between textbook availability and learning outcomes in cross-sectional data (see Piper et al. 2018 for a review), with similar examples using South African data (Gustafsson 2007; Spaull 2013). Yet these studies cannot identify the causal impact of textbooks because uptake and use are correlated with other factors such as motivation, prior skill, curriculum coverage, etc. More recent reviews of experimental research in education (McEwan 2015) include interventions with a materials component. However, they do not isolate materials from other components in an intervention and therefore cannot identify the impact of materials independent of other components (such as training). A more recent review of programmes to improve the quality of education in Africa included 145 studies and finds that the most successful programmes provide learner materials as part of a wider structured-pedagogy intervention (Evans & Acosta 2021), but the review does not isolate the impact of the materials specifically.

In that regard a number of international studies are worth highlighting. A well-known randomised control trial in Kenya found that distributing official government textbooks to learners had no impact on learning outcomes (Glewwe et al. 2009). These books were in English and were provided on a two-textbooks-to-three-children (67%)

1 The most common type of book to have been evaluated is the textbook, yet the word 'textbook' in this literature can mean any number of things and is often inclusive of dictionaries, workbooks, and readers in addition to traditional explanatory or instructional textbooks.

ratio. Subsequently it was found that the textbooks were too difficult for most learners and only benefited the top 20% of learners in the class (Glewwe et al. 2009). A similar study in Sierra Leone (Sabarwal et al. 2014, 1) also showed no impact of textbook provision but in this instance, it was because teachers hoarded the books: "A large majority of the books were stored rather than distributed to students…The propensity to save books was positively correlated with uncertainty on the part of head teachers regarding [future] government transfers of books". See also Falisse et al. (2020) showing positive effects of incentives to actually use textbooks provided in the Democratic Republic of Congo. We will return to this later in our chapter since in the present study the Eastern Cape Department of Education (ECDOE) seems to have convinced teachers that it would continue to provide the books each year, a promise it subsequently reneged on after two years.

The final international study to consider was a large, randomised control trial covering all rural schools in Mongolia. Fuje and Tandon (2018) evaluate the impact of classroom libraries and teacher training on reading outcomes. The three intervention arms were: (a) training only, (b) classroom libraries only, and (c) training and classroom libraries. This is in a context of low access to books or textbooks at school or at home. The classroom-library interventions included provision of 160 age-appropriate children's books[2] per classroom, as well as shelves to store them on. The authors found that "providing books had a strong impact on test scores. Books alone greatly increased test scores more than extra teacher training, but the books intervention still had a much weaker impact than training and books provided as a package" (Fuje & Tandon 2018, 1372). See Piper et al. (2018) for similar conclusions based on a study in Kenya testing the impacts of materials, training and coaching.

2.2 South African evidence of the impact of textbooks or books

Turning to South Africa, there are not many interventions that have rigorously evaluated the impact of a 'materials-only' intervention. Although there have been many large-scale interventions that are, in essence materials-only, few have been evaluated. Where studies are rigorously evaluated, it is rare to have a materials-only arm since the literature suggests this is unlikely to shift learning outcomes on its own. Studies evaluating Gauteng Primary Language and Mathematics Strategy (GPLMS) (Gauteng), Early Grade Reading Study (EGRS) (North West), and Funda Wande (Eastern Cape) all explicitly mention the importance of materials provision – and all of these interventions provided graded readers either in 'skinny-book' or anthology format – however, they were not designed to measure the impact of the materials independently from coaches or training or lesson plans (see Cilliers et al. 2019; Fleisch & Schoer 2014; Ardington & Meiring 2020; as well as Fleisch, this volume; Dornbrack & Kazungu 2022; Mtsatse, 2022). One additional study worth identifying is the 'Story Powered Schools' intervention – a randomised control trial evaluating the impact of a reading-for-

2 The cost of this intervention was $2.10 per book and $71.50 for a set of shelves (in 2008 prices) (Fuje & Tandon 2018,1363).

enjoyment programme that provided story books to rural schools. The programme was found to have no impact on reading outcomes or reading habits (Ardington et al. 2019). One recent randomised control trial (Funda Wande Limpopo) did include a materials-only arm in the intervention, and – usefully for our purposes – also included the Vula Bula anthologies of graded readers in a 1-1 ratio (see Makaluza & Mpeta, this volume). That intervention evaluated three intervention arms: (1) materials only, (2) materials and teacher assistants (1:1), and (3) control. After one year of intervention, the evaluation showed that there were no statistically significant differences between the materials-only arm and the control group (Ardington & Henry 2021, 6). In keeping with the findings from Mongolia, where materials were provided with additional support (in this case, a teacher assistant to every Grade 1 teacher) there was a large impact on reading outcomes (0.4 SD after one year). However, it should be noted that this was during the Covid-19 pandemic with rotational school timetables and significant disruption.

In summary, the existing literature suggests that providing 'books only' rarely has an impact on learning outcomes, although there are some exceptions where a large number of grade-appropriate books did improve reading outcomes (Fuje & Tandon 2018).

Although there is little evidence of materials-only interventions improving reading outcomes, it should also be noted that access to books (and textbooks in the Foundation Phase) is low. Data from the 2017 School Monitoring Survey showed that while 60% of Grade 6 learners had access to their own language textbook, at the Grade-3 level this was only 46% nationally and 40% in the Eastern Cape (DBE 2018, 46). Unfortunately, this lack of textbooks in Foundation Phase is not mitigated by school libraries or mobile libraries. The 2021 report of the National Education Infrastructure Management System (NEIMS) indicates that of 23,276 schools, 30% had libraries and only 17% had stocked libraries. The same report showed that the provinces with the lowest access to a stocked library were the Eastern Cape (4%) and Limpopo (3%).

2.3 Literature on basal readers

One particular type of 'textbook' is a reader where the stories included in it are sequenced and levelled moving from less difficult to more difficult. These are variously named graded readers, levelled readers, basal readers, and elsewhere "core reading programmes"[3]. The largest literature on readers comes from the United States (US). Reisboard and Jay (2013, 1) report that in the 20th century they were "the main instructional materials used in American elementary classrooms during the time". Similarly, Farr et al (1987, 265) report that "basal readers dominate reading instruction in most classrooms across the country. In fact, considering the central role of basal readers in the classroom, the selection of a basal reader represents the selection of a reading curriculum in most American schools". More recently, in a survey of 20,000 educators in 2017, as many as 65% of US educators used a basal programme for reading instruction (Simba Information 2017).

Almost all the research on graded readers internationally has been on English graded readers (see Hill 2012 for a review). This is problematic in a context such as South Africa

3 See Katz & Rees (this volume) and Dornbrack and Kazungu (2022) for explanations of the distinctions between these different types of readers.

where more than 70% of children are learning to read in an African language rather than English in the Foundation Phase (Spaull & Pretorius 2019). Furthermore, recent research on early grade reading programmes in Africa has also pointed to linguistic complexity as one of the explanatory factors of programme impact. In an impact evaluation of a mother-tongue reading-instruction programme implemented in 12 Ugandan languages, Brunette et al. (2019, 1) found that "language characteristics, including complexity, were more predictive of impact than implementation fidelity or socio-economic factors. We recommend that literacy improvement programme designs consider language complexity and characteristics, among other factors, to maximise impact". This is of special relevance to the present study given that the intervention materials (the Vula Bula anthologies) were specially designed for each target language, as we will explain in the next section.

2.4 The Vula Bula anthologies of graded readers

Katz and Rees (this volume) provide a detailed description of the history and development of the Vula Bula series in South Africa, and therefore those readers interested in the linguistic principles underlying their development are directed to their chapter in the present series. Very briefly, the books were developed as part of the GPLMS and funded by the Zenex Foundation. The development of the books was led by Jenny Katz at the Molteno Institute for Language and Literacies and started from the premise that pre-existing African language readers were "translated from English with little cognisance of isiXhosa phonic structures and little appreciation for the agglutinative nature of African languages" (Katz, 2019, i). The series has been independently evaluated by researchers from the South African Institute for Distance Education (SAIDE) who concluded that "the overriding impression we have of the materials is that they are based on sound educational and linguistic principles" (Sotuku & Kahla, 2013: 40). The series has also been endorsed by the minister of Basic Education: "the beauty of the Vula Bula Series is that…the whole series is available as Open Education Resources" (Motshekga 2015). This last element is important for two reasons related to cost: (1) Open Education Resources (OER) do not incur royalties or copyright costs, and (2) because the readers were OER, the Eastern Cape Department of Education (ECDOE) was able to collaborate with Molteno to combine the 'skinny books' into a single anthology per grade, both of which drastically reduce the costs of printing the books.

The 'traditional' form of graded readers are 'skinny' books of eight to 16 pages with each book containing a single story of the levelled series. According to the DBE's National Catalogue (DBE, 2021) the cost of one skinny book for a typical publisher was R16.10 (Platinum Grade R Series[4]). The cost of the Platinum Grade R Story Anthology (40 stories) was R499.80. If one were to purchase 22 Platinum Grade R readers it would cost R354. The total cost to the ECDOE for printing one Vula Bula anthology with 22 stories was R15. Put differently, the OER version cost 4% of the copyright version.

The small cost of the Vula Bula anthologies allowed the ECDOE to print one book per learner, which was given to the learner to own and keep. This is not the same for

4 We list the prices of the Grade R books only because the DBE's Catalogue was updated in 2021 for Grade R whereas the Grade 1-3 Catalogue was last updated in 2003 and thus the prices are out of date.

skinny books, which are kept by the teacher. Most teachers do not buy skinny-book readers on a 1:1 ratio for the class, but usually buy six to ten stories per level and then hand these out to learners depending on their reading level. Apart from the expense of this approach, there are also classroom-management and storage complexities with keeping 100+ skinny books ordered by level in a classroom, as well as greater demands on teacher judgement about which books should be allocated to which learners. In a context of low levels of literacy content knowledge among primary school teachers (Taylor & Mawoyo 2022; Murray et al. 2022), the anthologies require fewer professional judgements and are more logically and simply organised. The Vula Bula anthologies for Grades 1 and 2 were deliberately levelled from Story 1 to Story 44 such that if a learner struggles with Story 13, for example, they can move back to Story 12 or Story 11, etc.

In 2019, the first year of roll-out of the Vula Bula Anthologies, the ECDOE printed the grade-level anthology for all learners in Grades 1 to 3. In addition to their own grade-level anthology, learners in Grade 2 also received the Grade 1 anthology and learners in Grade 3 also received the Grade 2 anthology. The reason behind this was that local and international studies suggested that South African learners in the Foundation Phase were reading at levels one or two grades below curriculum expectations. In 2020, learners in all three grades received one reader each, with the expectation that they would keep their anthologies from the previous year.

2.5 The ECDOE roll-out of the Vula Bula anthologies

During 2018 the Eastern Cape Department of Education (ECDOE) developed its Reading Strategy for 2019-2023 (ECDOE, 2020). This included a section on the provision of Learner and Teacher Support Material (LTSM) and specifically graded readers. Early on in that process the Department considered open access readers due to the high cost of providing adequate numbers of graded readers in skinny-book format. Costs were driven by two factors (1) copyright and royalties paid to traditional publishers, and (2) the expense of printing books in skinny-book format – this is partly because the covers of books are more expensive than the inside pages (Vinjevold[5] 2018). As a result, the ECDOE partnered with a Non-Governmental Organisation, Funda Wande, which together with Molteno, was in the process of converting Vula Bula's open access graded readers into anthology format. By using open access materials and printing them in anthology format, the two largest cost-drivers were eliminated (royalties and book covers). This radically reduced the cost of printing the anthologies. Furthermore, by 'piggy-backing' off the existing delivery contract for the DBE Workbooks with a no-increase extension, it was possible to distribute all 825,000 anthologies to 4,298 schools with no additional delivery costs.[6]

In addition to the provision of the anthologies, the ECDOE arranged a five-day in-person training workshop for 200 subject advisors and lead teachers led by Funda

5 Penny Vinjevold was the Deputy Director General (DDG) Education Planning, Monitoring and Evaluation at the ECDOE from March 2017 to March 2020.

6 The printer and distribution agent in the Eastern Cape was Lebone Litho Printers who agreed to include the anthologies in the same plastic wrapping as the DBE Workbooks at no additional charge.

Wande literacy experts, together with ECDOE curriculum officials. This was conducted in October in 2018 (see Shalem et al., 2018) and again in June in 2019. In both instances the training was across four sites (Gqeberha, Komani/Queenstown, East London and Mthatha) with approximately 50 participants at each training facility.

In October 2018 the Superintendent General of the ECDOE, Mr Themba Kojana, sent out Circular 29 on "Learner and Teacher Support Material for 2019," which announced to all ECDOE schools that Foundation Phase learners would each be given anthologies of graded readers in January 2019 and that these should be taken home by learners (ECDOE 2018, 2).

3 Data

Data for this study are drawn from a large randomised controlled trial evaluation of Story Powered Schools (SPS), a reading-for-enjoyment programme, conducted in rural schools in two provinces in South Africa (Ardington et al. 2019). The staggered and clustered roll-out of the SPS programme necessitated conducting the evaluation over two cohorts of schools. Baseline-EGRA assessments were conducted with learners in Grades 2 to 4 in the first school term of 2017 and 2018 for cohorts 1 and 2 respectively. Endline EGRA assessments were conducted with the same learners in the third term of the following school year (2018 and 2019). Although the evaluation of that programme showed no differences in reading outcomes between treatment and control groups (Ardington et al. 2019) – that is to say there was no measurable impact of the intervention – the data collection in the control groups (where there was no SPS intervention) allows us to evaluate the impact of a concurrent intervention that was rolled out at the same time[7]. At the beginning of 2019 the Eastern Cape province provided the Vula Bula anthologies of graded readers to all Grade 1 to 3 learners in the province. The timing of the data collection and the roll-out of the anthologies allows us to measure the causal impact of these graded readers.

For our descriptive analysis of the delivery, uptake, and use of the anthologies, we use endline data from all 96 cohort 2 schools (SPS treatment and control-group schools) in the Eastern Cape. In total, we have 876 learner interviews, 168 teacher interviews, 93 principal or HOD interviews and 77 audits of Grade 2 classrooms.

To estimate the impact of the anthologies on learner reading outcomes, we restrict our focus to the 85 SPS control-group schools and use baseline and endline EGRA data to measure gains in reading skills. Thirty-six of these schools are from cohort 1 with the endline data collection completed before the roll-out of the Vula Bula anthologies. The remaining 49 schools are from cohort 2 with endline data collection happening three school terms after the roll-out of the Vula Bula anthologies. We have longitudinal data on 832 learners in the 36 cohort 1 schools and 1,250 learners in the 49 cohort 2 schools. Table 1 shows the sample size by cohort and grade and Figure 1 illustrates the comparison between cohorts.

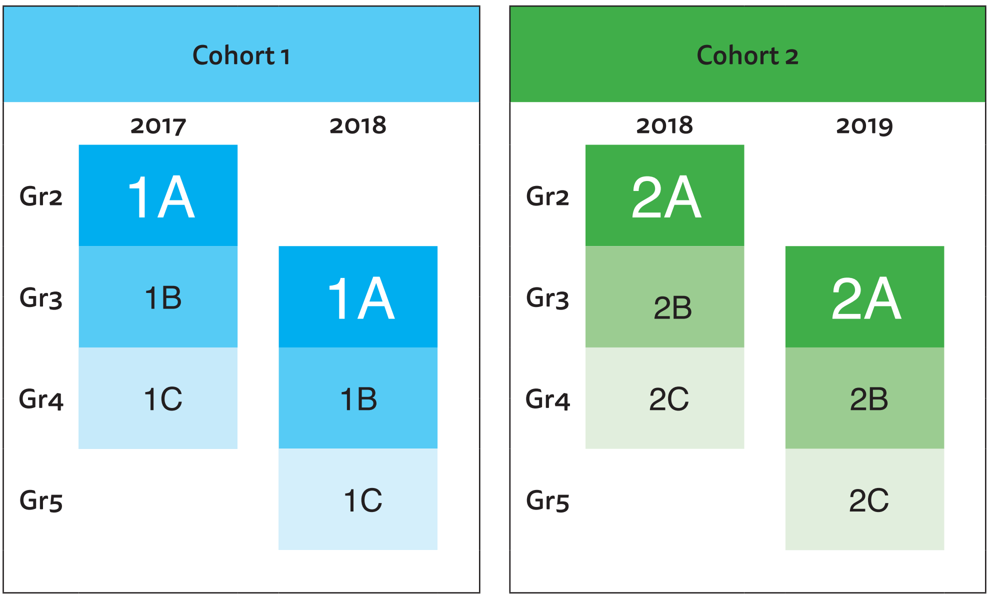
7 We do not include treatment schools as the SPS intervention included materials provision such as hanging libraries and cut-out-and-keep books.

Table 1 Sample size by cohort and grade

Baseline grade	Endline grade	Cohort 1: 36 schools Baseline: 2017 term 1 Endline: 2018 term 3	Cohort 2: 49 schools Baseline: 2018 term 1 Endline: 2019 term 3
Grade 2 term 1	Grade 3 term 3	Learners = 279	Learners = 384
Grade 3 term 1	Grade 4 term 3	Learners = 277	Learners = 382
Grade 4 term 1	Grade 5 term 3	Learners = 296	Learners = 366

Note: *We exclude the 7% of learners who were repeating the baseline grade at endline.*

Figure 1 Comparison of Cohort 1 and 2 by grade and testing dates (outlined square indicates cohort and grade that received the anthologies)



To identify the impact of the Vula Bula anthologies on learners' reading proficiency we exploit variation in exposure to the anthologies both within-grades-between-schools and between-grades-within-schools. Cohort 1 schools only received the Vula Bula anthologies in the year following the endline assessment while cohort 2 schools had the Vula Bula anthologies for three terms by the time of the endline assessment[8]. In cohort 2 schools, the Vula Bula anthologies were only distributed to the Grade 1 to 3 learners, and learners in higher grades should not have had access to the anthologies. Our treatment group who received Vula Bula anthologies is therefore the 384 learners

8 Contamination through schools from different cohorts sharing resources is not a concern here as all data collection in cohort 1 schools was completed before any Vula Bula anthologies were delivered. With respect to other graded readers, we are not aware of any other large scale provision that would have impacted the reading materials available in cohort 1 schools in 2017 and 2018. Indeed, at endline only 13% of learners in cohort one schools report that there are readers or story books in their classroom (Ardington et al. 2019).

who were in Grade 3 in cohort 2 schools at endline (cohort 2A in Figure 1). We will compare their reading gains against (a) Grade 3 learners in cohort 1 schools (cohort 1A) and (b) Grade 4 and 5 learners in cohort 2 schools (cohorts 2B and 2C) who would have been too old to benefit from the anthologies. This evaluation therefore measures the impact of the anthologies at the Grade 3 level only.

For each cohort, provincial officials selected the schools to be included in the evaluation. Thereafter, the evaluation team randomly assigned schools to treatment or control. This means that, while the treatment and control groups were balanced at baseline, there may be systematic differences in control schools across the two cohorts. We investigate this by comparing school and learner characteristics in cohort 1 and 2 control schools (Table 2). Learner characteristics are shown for the main comparison groups, those learners who were starting Grade 2 at baseline and in Grade 3 at endline.

Table 2 School and learner characteristics by cohort

	Cohort 1	Cohort 2
School characteristics		
Quintile 1	78%	76%
Quintile 2	22%	24%
Electricity	83%	94%
Flush toilet	3%	2%
Library	11%	20%
Grade 2 learner baseline characteristics		
No books other than schoolbooks to read at home	64%	77%
Learner is stunted (height-for-age z-score <-2)	11%	7%
Inside toilet	9%	4%
Refrigerator	57%	67%
Lives with mother	70%	74%
Lives with father	35%	33%
Correct letter sounds per minute (clspm)	23.8	19.3
Unable to read one word	28%	43%
Correct words per minute (cwpm)	8.5	6

For both cohort 1 and 2, around three-quarters of schools are classified as quintile 1 with the remainder being quintile 2 (i.e. these are the poorest of no-fee schools in the country). The vast majority of schools do have electricity but do not have flushing toilets (only one school per cohort (2-3%) has a flush toilet in the school). Learners have limited access to books at school and at home. Only one-in-ten cohort 1 schools and one-in-five cohort 2 schools have a central library or are visited by a mobile library. The majority of learners have no books other than schoolbooks to read at home (64% in cohort 1 schools and 77% in cohort 2 schools).

Across the two cohorts, learner characteristics are fairly similar although there are some differences, most notably the percentage of learners who are unable to read one word at the start of Grade 2. In the next section, we describe our methodology and discuss approaches to minimising any potential bias due to differences between the two cohorts.

4 Method

To estimate the impact of the Vula Bula anthologies, we use difference-in-differences (DD), a quasi-experimental technique. Our main DD estimate compares the reading gains of learners from the first term of Grade 2 to the third term of Grade 3 in cohort 2 schools against the gains of their peers in cohort 1 schools. With repeated observations on the same learner, we can write

$$Y_{ijct} = \alpha + \beta_1 C + \beta_2 T + \beta_3 (C \times T) + c_i + \varepsilon_{ijt},\ t = 1,2$$

where Y_{ijct} is the measured reading skill for student i in school j in cohort c at assessment point t. C is a dummy variable indicating whether the learner is in a cohort 2 school and T is another dummy variable that takes the value one for the endline assessment point (t=2). The term c_i is learner-level unobserved heterogeneity and ε_{ijt} is an independent and identically distributed error-term clustered at the school level to allow for correlation in the unobservables between learners within the same school. First differencing removes c_i and the DD estimate is

$$\beta_3 = (\overline{Y}_{c=2,t=2} - \overline{Y}_{c=2,t=1}) - (\overline{Y}_{c=1,t=2} - \overline{Y}_{c=1,t=1})$$

The DD assumes that, in the absence of the anthologies, the trends in reading skills would have been the same in cohort 1 and cohort 2 schools. While DD does not require the schools and learners to be identical across the two cohorts at baseline, the more inherently similar they are, the more plausible the assumption of parallel trends. We consider two approaches to improving the robustness of our DD estimates.

Firstly, we consider a triple difference (DDD) model where we strip out the difference in reading gains between Intermediate Phase (Grades 4 and 5 here) learners in cohort 1 and cohort 2 schools. This should remove any difference in trends between the two groups of schools that are not associated with the anthologies as these were only available to Foundation Phase (Grades R to 3) learners. We estimate the DDD by estimating ordinary least squares (OLS) regressions of the form

$$Y_{ijct} = \alpha + \beta_1 C + \beta_2 T + \beta_3 F + \beta_4 (C \times T) + \beta_5 (C \times F) + \beta_6 (T \times F) + \beta_7 (C \times T \times F) + c_i + \varepsilon_{ij},\ t = 1,2$$

where F is a dummy variable indicating whether the learner is in the Foundation Phase at endline. The DDD estimate is

$$\beta_7 = [(\overline{Y}_{c=2,g=3,t=2} - \overline{Y}_{c=2,g=3,t=1}) - (\overline{Y}_{c=1,g=3,t=2} - \overline{Y}_{c=1,g=3,t=1})] \\ - [(\overline{Y}_{c=2,g>3,t=2} - \overline{Y}_{c=2,g>3,t=1}) - (\overline{Y}_{c=1,g>3,t=2} - \overline{Y}_{c=1,g>3,t=1})]$$

where g is grade at endline. The term in the second set of square brackets is the DD for Intermediate Phase learners. To the extent that any differences in trends between the two groups of schools would be similar in the Foundation Phase and Intermediate Phase, subtracting this from the original DD removes such differences.

Our second approach to improving the robustness of our DD estimate is to use coarsened exact matching (CEM) to match the baseline-reading skills of cohort 2 learners with learners in cohort 1 schools. In addition to baseline-reading scores, we match on school quintile, indicators that the school has a library, the learner is classified as stunted, the presence of a refrigerator, inside toilet, and no schoolbooks to read in the home.

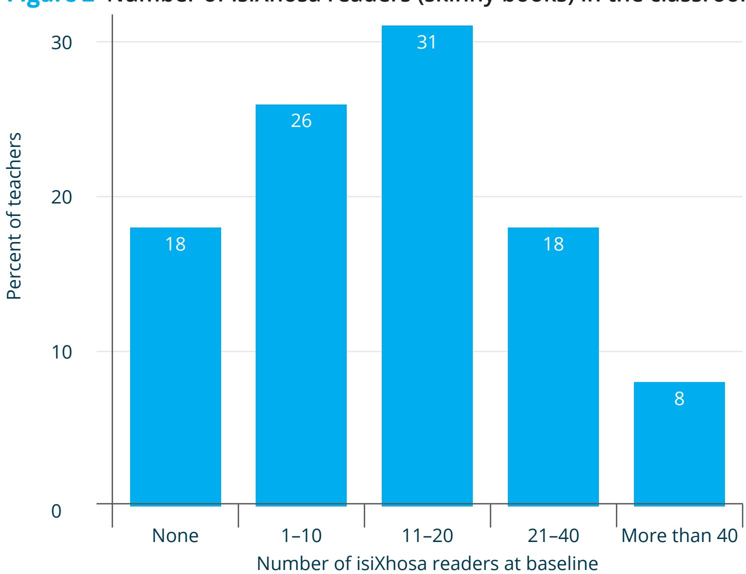
Before estimating the impact of the Vula Bula anthologies on reading outcomes, we consider the environment into which these anthologies were introduced and report on the delivery and use of these resources in cohort 2 schools.

5 Findings

5.1 Access to reading materials prior to distribution of anthologies

While libraries are fairly uncommon (Table 2), most teachers (82%) reported at baseline that they had at least one isiXhosa graded reader (a skinny book) in their classroom for the learners to read (Figure 2). However, the number of readers is woefully inadequate. When considering what an adequate or minimum number of skinny-book readers would be, it is important to understand that teachers typically order several copies of the same story and use these during group-guided reading. A number of curriculum experts would advocate for a minimum of one graded-reader story title per week, yielding 40 titles for one year (SAHRC 2021, 27). Figure 2 starkly illustrates just how book-scarce these classrooms were before the provision of the anthologies. The vast majority (92%) of teachers in the sample did not even have 40 readers in total, let alone 40 different stories. Furthermore, only about 21% of learners were ever allowed to take a reader home (as reported by the teacher).

Figure 2 Number of isiXhosa readers (skinny books) in the classroom at baseline



Note: *Number of readers includes multiple copies of the same title. Reports from 188 Foundation Phase teachers in all 96 cohort 2 schools at baseline*

5.2 Delivery and use

Almost all respondents (94%) to the principal/HOD interview reported that the Vula Bula anthologies were delivered to the school and in sufficient quantities for each learner to have their own copy[9]. Principals and HODs further reported almost universal distribution of anthologies to individual learners in classes. This is corroborated by the teachers, 82% of whom reported receiving enough copies for each learner to have their own copy (Table 3). Sixty-three per cent (63%) of Grade 3 learners reported that they were given their own copy of the Vula Bula anthologies. An additional 19% of Grade 3 learners reported that there are copies of the Vula Bula anthologies in their classroom. Both teachers and learners reported that the majority of learners are allowed to take the anthologies home.

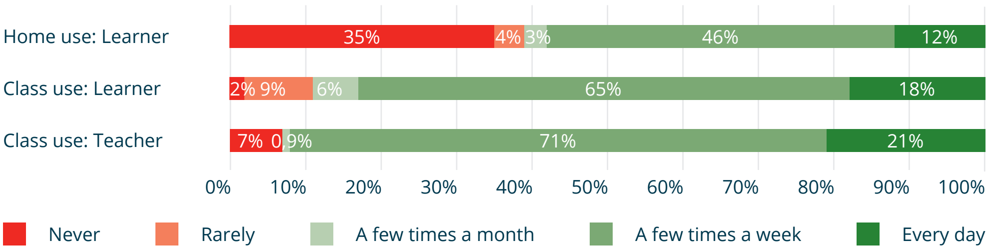
Table 3 Distribution of anthologies to learners

	Teacher report	Learner report
One per learner	82%	63%
Learners share	11%	19%
Learners allowed to take home	63%	55%

Notes: *Teacher reports from 168 teachers in 95 cohort 2 schools at endline. Learner reports from 876 learners in all 96 cohort 2 schools at endline.*

Figure 3 shows the frequency of reported use of the anthologies in class (teacher and learner reports) and at home (learner report). Around one in five teachers reported using the anthologies every day and 71% use them multiple times per week. The vast majority (83%) of learners reported using the anthologies at least once per week in class. Learners were also asked how often they read the Vula Bula anthologies at home and 58% reported reading them at least once per week at home.

Figure 3 Frequency of use of Vula Bula stories in class and at home

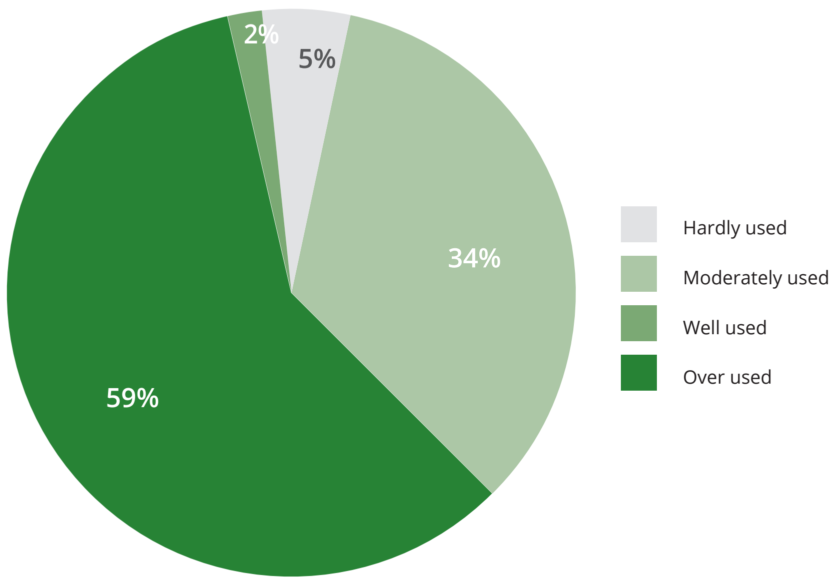


Note: *Teacher reports from 155 teachers across 90 schools as sample excludes teachers who report that they received no copies. Learner reports from 723 learners across 92 schools as sample excludes learners who report that they do not have their own copy nor are there copies available in the class.*

9 In Grade 3, 84% of principals report that there were sufficient copies delivered for each Grade 3 learner to have their own copy of the Grade 3 anthology but only 42% reported sufficient copies of the Grade 2 anthologies for each Grade 3 learner to receive their own copy.

During the visit to the Grade 2 classrooms in term 3, the enumerators asked the learners to show them their copy of the anthologies. In 55% of classrooms, at least three-quarters of learners were able to produce their own copy. The enumerators were asked to examine the anthologies and record how well-used they appeared to be (Figure 4). In 59% of the classrooms, the anthologies were considered well-used and in an additional 34% of classrooms they appeared to have been moderately used. Only 2% of classrooms had anthologies that were classified as so well-used that they were dirty, torn, or damaged.

Figure 4 **Condition of anthologies in Grade 2 classrooms – Classroom audit**



Note: *Sample is restricted to 61 of the 77 classrooms that were audited as learners were not able to show copies of the anthologies in 16 classrooms. Each classroom is from a different school.*

The delivery of the anthologies to schools and the distribution to learners was clearly effective. The teacher and learner reports, together with the classroom audit suggest extensive use of the Vula Bula anthologies in class and at home. In the next section, we examine whether the use of these resources resulted in improved reading outcomes.

5.3 Impact on learners' reading outcomes

Table 4 shows the estimated impact of the Vula Bula anthologies on learner reading outcomes, measured in correct words per minute. The main DD estimate shows that learners exposed to the anthologies (those in cohort 2 schools) improved their isiXhosa word-reading fluency by 3.2 words more than their peers who did not receive the anthologies (those in cohort 1 schools). Controlling for differences in the trends between learners in higher grades across the two groups of schools (DDD) reduces the estimated impact to 2.8 words per minute but it remains substantive and statistically significant.

Matching learners on baseline characteristics produces an estimated impact of 2.6 correct words per minute. The average gain in word-reading fluency in cohort 1 schools was 12.8 correct words per minute. The most conservative estimated impact of 2.6 words is equivalent to 0.19 standard deviations of cohort 1 learner gains and represents a 20% improvement over the gain in cohort 1 schools over the same period. A comparison with other Nguni-language early grade reading studies (Department of Basic Education 2019; Ardington and Meiring 2020) suggests that the provision of the anthologies resulted in improvements in fluency in the order of 16 to 21% over the normal rate of progression[10].

Table 4 Estimated impact in additional correct words per minute

	Control group gains in correct words per minute	Additional correct words per minute	Standard error of estimate	p-value
DD	12.8	3.17	1.18	0.009
DDD		2.79	1.21	0.024
DD with matching on selected baseline characteristics		2.55	1.25	0.044

The impact of the provision of anthologies is very much in line with that of other South African interventions aimed at improving early grade reading outcomes. The EGRS coaching programme improved the composite reading-proficiency score by 0.14 standard deviations in the first year and by 0.25 standard deviations over two years[11]. In the Eastern Cape at end of the first year in 2019, Funda Wande coaches increased composite-reading proficiency scores by 0.17 standard deviations.

In comparison to other interventions, the roll-out of the anthologies is very low-cost, making it extremely cost-effective. Assuming that it is sustainable to 'piggy-back' off the delivery of the DBE Workbooks at no additional cost, providing two anthologies per year would cost $2 per learner[12]. If we include training for all subject advisors and one head teacher per school we estimate a cost-per-learner of around $9[13]. As points of comparison, the EGRS coaching arm and the SPS programme were delivered at a cost of around $42 per learner.

10 The estimated impact was compared to the gains for control group learners from 1) the end of Grade 2 to the end of Grade 3 for the first Early Grade Reading Study (EGRS I) and 2) the beginning of Grade 2 to the end of Grade 2 for the Funda Wande evaluation in the Eastern Cape. Gains were scaled by the number of terms between the two evaluation points.

11 On the word-reading task, treatment learners were reading, on average, 3.8 correct words per minute faster than their peers in control schools (Cilliers et al. 2019).

12 In 2019, learners in Grades 2 and 3 received two readers at a cost of R15 ($1) per reader.

13 The provincial department arranged a 5-day in-person training workshop for 200 subject advisers and lead teachers led by Funda Wande literacy experts together with curriculum officials. In the absence of cost data for the training, we use the costs associated with teacher training for the SPS project as an estimate. The SPS two-day training for teachers cost $298 per teacher (Ardington et al. 2019). Scaling up to five days and budgeting for one lead teacher from each of the 4,298 schools and 200 subject advisors, the cost per learner for training is around $7.

6 Conclusion

In this chapter we report the findings of an evaluation of the provincial roll-out of a government-funded initiative to provide anthologies of graded readers to all Grade 1, 2, and 3 learners in the province. This was accomplished by (1) identifying high-quality Open Access readers in the home language of the learners, (2) compiling them into a single anthology per grade, and (3) streamlining delivery by using existing contracts and distribution networks. This allowed the ECDOE to commission the printing and delivery of 825,000 anthologies to 463,276 children across 4,298 schools at a cost of approximately R15 ($1) per anthology.

The independent evaluation showed a high level of intervention fidelity with anthologies being delivered to schools in line with delivery records and the provincial department's audit. The anthologies were distributed to learners, with 63% reporting that they were given their own copy. Crucially, multiple sources of evidence indicate that the anthologies were very well used and not only in the classroom. The majority of learners report reading the anthologies at least once a week at home.

Our difference-in-difference estimates showed that learners who were exposed to the anthologies could read three additional isiXhosa words correctly per minute in Grade 3 compared to those learners who were not. This finding is robust when comparing across schools within the same grade and across grades in the same school.

Our findings are consistent with more recent research on textbook provision where usage was incentivised (Falisse et al. 2020). Although earlier international research found no or selective impacts, usage of the textbooks was very low (Glewwe et al. 2009; Sabarwal et al. 2014). In South Africa, the materials-only arm of the Funda Wande intervention in Limpopo was ineffective but the school year was seriously disrupted by Covid, limiting learner access to materials. Nevertheless, it is surprising to see impacts for the low-cost provision of anthologies in line with multi-faceted structured pedagogy programmes that include materials provision, training, and ongoing coaching support. We consider some of the factors that may have contributed to both the high usage and positive impact of the programme. First, the anthologies were provided in a deep rural context where learners had extremely limited access to home-language reading materials. Quantities of readers were woefully inadequate and only 6% of learners in cohort 1 schools reported that they were allowed to take school readers or story books home. Second, the provision of the anthologies in a 1-1 ratio with explicit instructions from the department that learners be allowed to take them home, together with expectations that they would be provided each year, appears to have been very effective in overcoming potential hoarding of these materials. Third, the Vula Bula anthologies were rigorously developed with consideration of the orthographic features of each language and how these impact learning to read. Extensive usage in class suggests a high degree of acceptability from the teachers' perspective.

It is not possible to disentangle which element caused this improvement. Was it the better teaching of reading due to more and better-graded readers? Or was it the fact that children owned the books themselves and could take them home? Or was it that they were distributed on a 1-1 ratio as part of a formal province-wide government initiative? Or was it that they were phonically-sequenced, logically-graded and simple to use? It is likely that it is a combination of all of these.

From a policy point of view, we think the results are important for several reasons:

1. They show that it is possible to implement innovative low-cost interventions at scale that lead to modest improvements in reading outcomes, at least when the baseline level of resources is very low, as it is in many parts of South Africa and the rest of the continent.
2. They illustrate the potential of open access materials being combined in innovative and pedagogically helpful ways to support teaching and learning, and – critically – to significantly reduce the cost of material provision.

On a cautionary note, it must be stated that while the programme was rolled out for two years (2019 and 2020), in 2021 and 2022 the books were not printed or distributed despite being an effective intervention and formally being part of the ECDOE Reading Plan (ECDOE, 2020:2). The exact reasons are unknown but possibly relate to budgetary pressures and the prioritisation of digital tablets for Grade 12 learners in 2020. This yo-yo effect of new interventions (providing but then withdrawing) may help explain why teachers often hoard new materials and do not give them out to learners (see Sabarwal et al, 2014).

This chapter provides some of the first research findings on African-language graded readers and establishes the effectiveness of a low-cost scalable model for addressing the lack of appropriate African-language texts to teach children to read in South African classrooms.

References

Ardington, C. and Henry, J. (2021). Impact Evaluation of Funda Wande Teacher Assistant Intervention in Limpopo Province South African Labour Development Research Unit (SALDRU), University of Cape Town. Cape Town: SALDRU

Ardington, C. and Meiring, T. (2020). Impact Evaluation of Funda Wande Coaching Intervention Midline Findings South African Labour Development Research Unit (SALDRU), University of Cape Town. Cape Town: SALDRU

Ardington, C., Hoadley, U. and A. Menendez (2019). Impact Evaluation of USAID/South Africa Story Powered School Program Endline Report. NORC at the University of Chicago. Washington: USAID.

Brunette, T., Piper, B., Jordan, R., King, S. and Nabacwa, R. (2019). The Impact of Mother Tongue Reading Instruction in Twelve Ugandan Languages and the Role of Language Complexity, Socioeconomic Factors, and Program Implementation. *Comparative Education Review*, 63(4) 591–612.

Cilliers, J., Fleisch, B., Prinsloo, C. and Taylor, S. (2019). How to Improve Teaching Practice?: An Experimental Comparison of Centralized Training and In-Classroom Coaching. *The Journal of Human Resources*, 55, 926–962.

Department of Basic Education. (2018). School Monitoring Survey 2017/2018: Technical Report. Pretoria: Department of Basic Education.

Department of Basic Education. (2019). The Secondary Early Grade Reading Study. Year 1 Report. Learner performance after the first year of implementation. Pretoria: Department of Basic Education.

Department of Basic Education. (2020). Action Plan to 2024: Towards the Realisation of Schooling 2030. Pretoria: Department of Basic Education.

Department of Basic Education. (2021). National Education Infrastructure Management System Report. Pretoria: Department of Basic Education.

DPME (Department of Planning, Monitoring and Evaluation). (2020). Medium Term Strategic Framework: 2019-2024.

Dornbrack, J. and Kazungu, T (2022). Foundational skills in home language reading in South Africa: Empirical evidence from 2015–2021 In Spaull, N. & Pretorius, E. (Eds) *Early Grade Reading in South Africa*. Cape Town: Oxford University Press.

ECDOE. (2018). Department Circular No.29 of 2018 LTSM for 2019. Eastern Cape Department of Education.

Evans, D. K. and Acosta, A. M. (2021). Education in Africa: What Are We Learning? *Journal of African Economies*. 30(1), 13–54.

Falisse, J.B., Huysentruyt, M. and Olofsgård, A. 2020. Incentivizing Textbooks for Self-Study: Experimental Evidence from the Democratic Republic of Congo. Academy of Management Proceedings 1, 19578.

Farr, R., Tulley, M. A. and Powell, D. (1987). The Evaluation and Selection of Basal Readers. *The Elementary School Journal*. 87(3), 266–281.

Fleisch, B. and 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), 1–12.

Fuje H. and Tandon, P. (2018). When do in-service teacher training and books improve student achievement? Experimental evidence from Mongolia. *Review of Development Economics*. 22(3), 1360–1383.

Glewwe, P., Kremer, M. and Moulin, S. (2009). Many Children Left Behind? Textbooks and Test Scores in Kenya. *American Economic Journal: Applied Economics*, 1(1), 112–135.

Graham, J. and Kelly, S. (2018). How Effective are Early Grade Reading Interventions? A Review of the Evidence (English). Policy Research Working Paper, no. WPS 8292 Washington, D.C.: World Bank Group.

Gustafsson, M. (2007). Using the Hierarchical Linear Model to Understand School Production in South Africa. *South African Journal of Economics*, 75(1), 84–98.

Hill, D. (2012). Graded Readers. *ELT Journal*, 67(1), 85–125.

Howie, S. J., Combrinck, C. M., Roux, K., Palane, M. and Mokoena, M. (2017). South Africa Grade 4 PIRLS Literacy 2016 Highlights Report: South Africa.

Katz, J. L. (2019). Investigating the Linguistic Effectiveness of Early Reading Schemes in isiXhosa: A Phonological and Orthographical Analysis of Three isiXhosa Grade 1 Graded Reader Series. Rhodes University. Masters Thesis. Available at: http://hdl.handle.net/10962/92798 (Accessed 21 February 2022)

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.

Molteno. (2018). isiXhosa Incwadi yamabili Ibanga lesi-3. Available at: https://fundawande.org/img/cms/resources/Xho%20Booklet%20Gr3.pdf (Accessed 10 March 2022)

Moore, A., Gove, A. and Tietjen, K. (2017). Great Expectations: Key Factors Affecting Student Learning of Foundational Reading Skills. *New Directions for Child and Adolescent Development*. (155), 13–30.

Motshekga, A. (2015). Minister Angie Motshekga: Zenex Foundation 20th Years Celebratory Event. (Online). 22 October, 2015. Available at: https://www.gov.za/speeches/minister-angie-motshekga-zenex-foundation-20th-years-celebratory-event-22-oct-2015-0000 (Accessed 21 February 2022)

Murray, S., Taylor, N., & Nkomo, A. (2022). In-service Teacher Education for Foundation Phase Literacy: The case of the Advanced Certificate in Foundation Phase Literacy at Rhodes University In Spaull, N. & Pretorius, E. (Eds) *Early Grade Reading in South Africa*. Cape Town: Oxford University Press.

Piper, B., Zuilkowski, S. S., Dubeck, M., Jepkemei, E. and Simon, J. K. (2018). Identifying the Essential Ingredients to Literacy and Numeracy Improvement: Teacher Professional Development and Coaching, Student Textbooks, and Structured Teachers' Guides. *World Development*. 106, 324–336.

Read, T. (2015). *Where Have All the Textbooks Gone?: Toward Sustainable Provision of Teaching and Learning Materials in Sub-Saharan Africa*. World Bank Group.

Reisboard, D. and Jay, A. B. (2013). Teachers' Perceptions: Transitioning from Teacher Selected Reading Materials to Implementing a Core Reading Program. *The Journal of Balanced Literacy Research and Instruction*, 1(2), 6.

Sabarwal, S., Evans, D. K. and Marshak, A. (2014). *The Permanent Input Hypothesis: The Case of Textbooks and (No) Student Learning in Sierra Leone*. Policy Research Working Paper: No. 7021. World Bank Group, Washington, DC.

Shalem, Y., Dawber, A., & Koornhof, H. (2018). Funda Wande Teaching Reading in The Foundation Phase: Evaluation Report of ECDOE Training 1-5 October 2018. University of Witwatersrand. Johannesburg.

Simba Information. (2017). Are New Standards Changing Reading Instruction and Curriculum Selection? Online. Available at: https://www.simbainformation.com/Content/Blog/2017/05/17/Are-New-Standards-Changing-Reading-Instruction-and-Curriculum-Selection (Accessed 10 February 2022).

Sotuku, N., and Kahla, L. (2013). Final Report on Materials Evaluation: Molteno Vula Bula isiXhosa Level 1 Language and Literacy. South African Institute for Distance Education (SAIDE). Pretoria: SAIDE.

Spaull, N. (2013). Poverty & privilege: Primary School Inequality in South Africa. *International Journal of Educational Development*, 33(5), 436–447.

Spaull, N. (2022). 2022 Background Report for the 2030 Reading Panel. Cape Town. Available at: https://www.readingpanel.co.za/_files/ugd/e4cf67_fc97218559fb49dc9afd7942a81a4f61.pdf (Accessed 21 February 2022).

Spaull, N. and Pretorius, E. (2019). Still Falling at the First Hurdle: Early Grade Reading Outcomes in South Africa. In Spaull, N. & Jansen, J. (Eds): *South African Schooling: The Enigma of Inequality*. Cham: Springer.

Stern, J., Jukes, M., Piper, B., DeStefano, J., Mejia, J., Dubeck, P., Carrol, B., Jordan, R Gatuyu, C., Nduku, T., Punjabi, M., Harris Van Keuren, C. and Tufail, F. (2021). Learning at Scale: Interim Report. RTI International.

Taylor, N. & Mawoyo, M. (2022). Professionalising teaching: The case of language and literacy. In Spaull, N. & Pretorius, E. (Eds) *Early Grade Reading in South Africa*. Cape Town: Oxford University Press.

UNESCO. (2016). Every Child Should Have a Textbook. Global Education Monitoring Report. Policy Paper 23. Available at: https://unesdoc.unesco.org/ark:/48223/pf0000243321 (Accessed 21 February 2022).

Vinjevold, P. (2018). Personal Communication

08

Two birds with one stone: Improving Grade 1 learning outcomes using Funda Wande teaching assistants in Limpopo 2021-2022

NWABISA MAKALUZA & BOKANG MPETA

Abstract

This chapter presents a discussion about the design and implementation of the Funda Wande Limpopo Teaching Assistant structured-pedagogy intervention in 2021 and 2022. The randomised control trial showed that after one year (2021), the teaching assistant intervention achieved statistically significant improvements in both literacy and mathematics outcomes at the Grade-1 level, with learners performing 0.44 standard deviations higher in literacy and 0.38 standard deviations higher in numeracy when compared to control schools. These are the largest effect-sizes seen in South African primary school interventions to date. The four main features of the TA intervention were: (1) the rigorous recruitment and selection of youth, (2) the provision of high-quality learner- and teacher-support material, (3) training that is both comprehensive and aligned to the materials provided, and (4) continuous monitoring and support. We also reflect on what worked well and what did not work well. The intervention occurred when South Africa was hit by the Covid-19 pandemic with concomitant learning losses. It is for this reason that we argue that teaching assistants, when properly recruited, trained, monitored, and supported, have the potential to play an important role in a scaled-up learning recovery programme.

KEYWORDS
teaching assistant, youth unemployment, learning losses, early grade learning, Foundation Phase, Home Language, Literacy, Numeracy, Coronavirus, Covid-19, structured pedagogy

1 Introduction

South Africa's ongoing early-learning crisis has only been exacerbated by the pandemic. The 2019 Trends in International Mathematics and Science Study (TIMSS) reports that 63% of Grade 5 learners could not reach the low international benchmark in the less difficult mathematics assessment set at a Grade 4 level (Reddy et al. 2020). The widely-cited Progress in International Reading and Literacy Study (PIRLS) in 2016 showed that 78% of learners in South Africa could not reach the low international reading-and-literacy benchmark at the end of Grade 4 (Howie et al. 2017, see also Van Staden & Gustafsson 2022). This meant that 78% of South African learners could not read with understanding in the language of assessment, which was also the Language of Learning and Teaching (LOLT) in the Foundation Phase. Of all 11 languages, performance was the poorest in Sepedi, a language spoken predominantly in the Limpopo province of South Africa. Amongst the recommendations offered in the report were reducing class sizes, increasing the share of time spent on reading activities, attracting young people to the teaching profession to mitigate teacher attrition, and focusing interventions on at-risk groups within the system (Howie et al. 2017).

Teaching Assistants (TAs) have been used in many contexts to provide additional capacity in the classroom, which allows students to receive more focused attention. In South Africa, it is widely known that quintiles one to three or so-called 'no-fee' schools have significant resource constraints and thus face issues such as overcrowding, which increase barriers to learning. In our 120 no-fee-school sample, for example, the average Grade 1 class size in 2021 was 41 learners per class. Employing TAs in these under-resourced, overcrowded schools can ameliorate some of the constraints educators face, giving them more space to focus on the business of teaching, and creating more opportunities for learners to get the support they need. This is the hypothesis being tested in the 2021-2022 Funda Wande Limpopo Teaching Assistant Intervention.

Funda Wande is a not-for-profit organisation whose mission is to ensure that all children in South Africa can read for meaning and calculate with confidence by the age of 10. The TA intervention is a randomised control trial (RCT) with 120 schools in total in the Capricorn North and Capricorn South districts of Limpopo. Schools were randomly allocated to one of three arms, with 40 schools in each arm: (1) TA and learning and teacher-support material (LTSM), (2) LTSM and centralised training, and (3) control.

The external midline evaluation of the Funda Wande Limpopo TA intervention[1] found learners in the 'TA and LTSM[2]' arm of the RCT to be at a distinct advantage compared to learners in the control arm. Learners in schools that had Funda Wande TAs performed better on a range of literacy and mathematics tasks, scoring higher on composite measures for the Early Grade Reading Assessment (EGRA) and Early Grade Mathematics Assessments (EGMA) by 0.44 and 0.38 standard deviations, respectively (Ardington and Henry, 2021). Juxtaposing this with estimates that Foundation Phase learners in the country experienced learning losses of up to 60% of a normal school year[3] due to the Coronavirus pandemic, it is evident that the TAs in the intervention

1 Data for midline evaluation was collected during term 3 and the first part of term 4, between 19 August and 20 October 2021.

2 Hereafter referred to as the TA arm

3 Learning losses on four EGRA sub-tasks ranged between 0.26 and 0.43 standard deviations.

played a vital role in buffering their learners from Covid-19-related learning losses during 2021 (Ardington et al., 2021).

This chapter describes how these results were achieved by answering the following research question: what type of programme can get unemployed youth with a minimum of Grade 12 to improve early-learning outcomes in mathematics and home language (literacy) in no-fee schools in Grade 1?

We begin by outlining some of the literature on TAs, followed by an overview of the design and implementation of the TA intervention. This entails a description of the school selection, TA roles and recruitment, and the structured pedagogy approach to implementation. Thereafter we discuss how this programme speaks to the results of the external evaluation by Ardington and Henry (2021). Lastly, we discuss the learnings and opportunities for growth within the TA intervention's scope and as they relate to the education system more broadly. As the two programme leads for the Intervention, we believe we are well-placed to document the challenges and successes of this recent intervention.

2 Literature and background

2.1 The Funda Wande TA intervention

It is clear that most learners in the South African education system do not reach their grade-level proficiency. It is therefore important to set a level of instruction that encourages learners to meaningfully engage in the learning of content, i.e. teaching children at their learning level rather than where they are 'meant' to be according to the curriculum. This view was critical in defining the roles and responsibilities of TAs within the Funda Wande Limpopo intervention.

TAs were instructed to help the teacher identify struggling learners and work with them individually or in small groups of approximately five to six learners. This method was effective in Ghana (Duflo et al. 2020) and India (Banerjee et al. 2007). Additionally, high-quality structured resources were provided to the teachers and learners (Mtsatse 2022; Sapire et al. this volume).TAs were trained on each term's content using these same resources. The training was conducted by TA mentors, usually at the start of the school term. TAs were then supported over the year by the mentors, who themselves were trained by Funda Wande content specialists. Support for this approach was in a study done by Pratham that found that having an additional layer of supervisors improved learning outcomes (Banerjee et al. 2016).

2.2 Learning losses induced by lockdown

The current intervention was implemented in the most unstable time in recent history: the Coronavirus pandemic. Schools in South Africa were required to reduce the number of learners in classes by using rotational timetabling[4], which meant that learners

4 Learners in a class are separated into two or three groups and these groups take turns to attend schools in rotation.

had 118 days at school of the 198 days allocated in the school calendar (Shepherd & Mohohlwane 2021).

Early research on the impact of Covid-19-induced lockdowns showed that children in the most vulnerable parts of the schooling system lost about 60% of the learning outcomes of a normal school year (Ardington et al. 2021). This outcome was a major setback in working toward the shared national goal announced by the president in 2019: "within the next decade, every 10-year-old will be able to read for meaning" (SONA 2019). Most of these large learning losses can be attributed to children not having enough instructional time due to the rotational timetables and school closures.

We have no reason to believe that the learning losses for early grade mathematics are any less severe. The pre-pandemic 2019 TIMSS-Numeracy study showed that 63% of learners in Grade 5 could not add, subtract, multiply, or divide one- or two-digit numbers, nor could they solve simple word problems (Reddy et al. 2020; Spaull et al. 2022). These basic mathematical skills should have been acquired in the Foundation Phase.

The form of the rotational timetable was left to the schools' discretion. Consequently, learners had limited and uneven exposure to the curriculum and the Funda Wande programme. Funda Wande also had to iteratively experiment with different training modes (virtual, in-person, or hybrid) as a response to national surges in Covid-19 cases. Based on learnings and feedback from the intervention team, the implementers evaluated and adapted to the ever-changing environment brought on by the Covid-19 pandemic.

Rotational timetables were ended at the beginning of 2022. Though encouraging, teachers had inadvertently grown accustomed to the new normal of smaller class sizes (typically 50% of their 'normal' class sizes) during the extended implementation of rotational timetables. Reintroducing full class sizes of roughly 65 learners per teacher may have left teachers feeling overwhelmed. Teachers have listed large class sizes as one of the top priorities for policy changes (OECD, 2020). New evidence of increasing large (50+) and extreme (60+) class sizes between 2015 and 2019 are also cause for concern (Reddy et al. 2020; Spaull et al. 2022). Within this context, TAs can be employed in a structured programme to help teachers manage large class sizes.

3 Intervention design and implementation

3.1 School selection

In collaboration with the Limpopo Department of Education (LPDOE), Funda Wande recruited 120 primary schools in the Capricorn North and South districts. To qualify, schools had to meet the following criteria:

- No-fee (quintile 1-3)
- Sepedi language of learning and teaching (LOLT) in Foundation Phase
- Not have chronic management issues
- The learner-educator ratio should not exceed 65 learners per educator, nor have fewer than 20 learners per educator
- Not have more than four classes per grade
- No multigrade classes
- Not part of an existing project for early-learning outcomes in the Foundation Phase.

The broad rationale for the above criteria was to identify 'typical' no-fee schools. To that end, we wanted to exclude very large and very small schools and those completely dysfunctional (i.e., in an ongoing dispute with the Department). We also did not want to confuse teachers by introducing a new programme when they were already trying to implement another. From the 2019 EMIS master list, of the 519 no-fee primary schools in the Capricorn North and South districts in Limpopo, 379 (73%) were eligible for applications. Funda Wande invited these schools to apply for the intervention and received 170 applications which were vetted. We selected 120 schools for randomisation. This selection pre-randomisation was based on the motivation written by the principal and Foundation Phase head of department (HOD) in the application letter. The 120 schools were randomly allocated into three arms of 40 schools each: (1) TA arm, (2) LTSM arm, and (3) control arm. The programme supported Grades R and 1 in 2021 and phased in Grade 2 in 2022. All intervention schools received a comprehensive set of materials from the literacy (see Mtsatse 2022) and mathematics (see Sapire et al., this volume) programmes. These materials included Learner Workbooks and Teacher Guides (one for language and life skills, and one for mathematics), which structure and sequence the entire curriculum. Additional resources such as Vula Bula anthologies on a 1-1 ratio (see Katz & Rees, this volume) and mathematics manipulatives were provided for learners in both intervention arms. In addition, all teachers in the intervention schools received quarterly training, and any additional resources (e.g. modest classroom libraries[5]) were provided to all three arms to maintain balance in the sample.

3.2 Recruitment and roles of teaching assistants

3.2.1 Recruitment of teaching assistants

Sourcing and filtering

Recruitment of TAs followed the allocation of schools to the three arms. Funda Wande had to answer an important question: how to best recruit from a large pool of unemployed youth in a short timeframe. Recruitment involved a three-stage process: sourcing and filtering applicants, conducting competency assessments, and a final selection day.

The opportunity was advertised at the TA intervention schools and other hotspots that would attract youth. Interested youth were required to register online so that they could be filtered according to basic criteria. Applicants needed to be unemployed at the time of recruitment, between the ages of 18 and 28, fluent Sepedi speakers, with at least a Grade 12 qualification (or NQF level 4 equivalent). It was also important to recruit TAs from the communities that the schools served as this would limit churning caused by poor fit and high transport costs. Youth who met these criteria were invited to take a set of competency assessments to evaluate their suitability for the programme.

5 Breadline supplied 21 classroom libraries across all three arms of the intervention. Each classroom library had 78 titles from Biblionef, Book Dash, Mikhulu Trust, and Bargain Books.

Competency assessments

Assessments were online, and applicants would need at least an entry-level smartphone and internet connectivity to access them. The competency assessments could be completed in less than ninety minutes. Six assessment hubs were set up at schools across the intervention districts equipped with internet connectivity and devices to loan. The absence of these hubs and devices would have excluded many prospective candidates because of their economic vulnerability and not their competencies. Roughly 1,200 applicants were assessed over three days[6].

The competency assessment was divided into sections, which tested applicants' numerical and language reasoning, work values, preference profile, and personal attributes and skills. The numeracy assessment used items at a similar level to released items from the TIMSS Grades 5 and 9. The language reasoning assessment was conducted in Sepedi and tested their comprehension skills and fluency. The work values, preference profile, and personal attributes and skills sections included basic self-reported questions that would help to rule out candidates who would be very obviously unsuited to a teaching-assistant role. A minimum score of 50% on all the sections was required to pass the assessment. This rule was only relaxed where not enough candidates at a particular school had passed[7]. Although all the assessed candidates had passed Grade 12 or NQF 4, less than 40% of the assessed candidates passed the numeracy and literacy assessments. The highest-scoring 256 youth were invited to the final round.

Final selection

The final round of the recruitment process was a selection day bootcamp where eight Funda Wande staff could meet the 256 candidates and observe their interpersonal, problem-solving and communication skills, and ability to work in teams. The selection day had four one-hour-long stations: (1) literacy, (2) mathematics, (3) Harambee Youth Employment Accelerator's Choose-to-win[8], and (4) an administration station[9]. Thus, each candidate spent four hours at selection, rotating from one station to the next in groups of 16. This process was repeated until all 256 candidates had been observed.

The best-performing 145 candidates were appointed as TAs, 77 of whom were allocated to Grade 1, the focus of this chapter. The TAs' salaries were funded by the Youth Employment Service (YES), which paid a minimum wage (R3,760 per month in 2021). The TAs were then appointed and trained before their first entry into schools.

---

6 A special arrangement was made for applicants from schools in very remote areas who could not make it to the hubs. This extended the finalisation of the recruitment process by two weeks.

7 There were some schools where not enough applicants had passed all sections of the competency assessments. In such cases we allowed applicants to move onto the next round if they obtained 50% for the Numerical and Language Reasoning sections.

8 A short course to observe the candidates' soft skills such as punctuality, positive attitude, and curiosity. This was a series of videos followed by prompting questions for small group discussions.

9 Candidates submitted certified copies of their identity document, Grade 12 certificate, a signed affidavit stating that they were unemployed at the time of recruitment, proof of residence, and completed and signed MIE indemnity form. Their fingerprints were captured for criminal background checks. We submitted the above documents to MIE (mie.co.za) which did the verification checks within two business days.

3.2.2 The role of teaching assistants

Schools selected for the TA intervention received one TA per teacher. The TAs attended school daily and mostly remained with their assigned teacher for the year. Funda Wande recruited six TA mentors who were responsible for oversight and monitoring the 145 TAs. Each mentor was responsible for 21 to 26 TAs which translated into one mentor overseeing four to ten schools, depending on the number of classes per grade. A further description of the mentors is in Section 3.3. below. These mentors facilitated the introduction of TAs to schools at the beginning of the year. They were required to have meetings with the TAs, their teachers and, if possible, the school principal and/or HOD. The mentors also had to discuss the partnership agreement and create a shared understanding of the relationship between the TA, Funda Wande and the school. The partnership agreement detailed the duties, responsibilities, and limitations of the TAs' role, as well as some guidelines on how the teachers and TAs should relate to each other, which are discussed below.

TAs could be expected to perform the following duties:

- Assisting the teacher during class for the duration of the school day. TAs were not meant to be general school assistants or cleaning staff but were expected to create an environment that would facilitate improved teacher practice.
- Assisting their teacher with administration (such as taking roll call or marking).
- Helping the teacher set up the classroom before lessons and pack away materials afterwards.
- Helping the teacher identify and support learners who were struggling to grasp certain skills and concepts.
- Helping the teacher oversee and assist the learners during whole-class and independent activities.
- Revising concepts with learners that teachers had previously taught.
- Reading aloud to learners, assisting learners with independent reading, handwriting, and other tasks for which the TA has been trained.
- Doing remedial exercises and games with learners during class or break times.
- The TA was also obligated to report to their mentor – who would, in turn, report to the school management team and the Funda Wande TA project manager – any incident where the rights of any children as outlined in the Children's Act. No. 38 of 2005 were violated, for example, corporal punishment and any abuse of a sexual or physical nature.

TAs were not trained for nor permitted to perform the following tasks:

- Teaching new material/curriculum to learners in small groups or the whole class.
- Ensuring that the curriculum was completed.
- Covering any financial costs of learners or teaching.
- Arranging and/or supervising extracurricular trips or activities for learners.

- Communicating with parents regarding the progress of or any challenges learners face.
- Supervising classes alone for more than 15 minutes[10].

To create a healthy working environment between the teacher, school management, and TA, some parameters and conduct guidelines had to be established, understood, and mutually agreed upon. Both teacher and TA had to foster a relationship of mutual respect where the teacher would guide and mentor the TA. Should disagreements arise between the two parties, both ought to try and reach a resolution before involving a third party (ideally a more senior member of the school's management and/or the mentor). It was important for the TA to be punctual and inform the teacher should anything delay them or necessitate them being absent. The TA was to remain in the school until the teacher indicated that they may leave.

3.2.3 TA monitoring and support

Mentors were responsible for the training, monitoring, and support of TAs. The selection of mentors was done in three phases: two mentors were hired in September 2020, three were hired in October 2020, and one in March 2021. The mentors had tertiary qualifications in education, linguistics, social work, and psychology, among others. They also all had some experience working with learners in some capacity. They had a strong affinity for and prior experience working with youth. Although advantageous, they were not required to have a professional qualification in education and/or experience working in the Foundation Phase. Rather, they were required to have the attributes and experience that would make them effective trainers and good mentors to young people. Of the six mentors, one was male, and five were female.

As far as possible, schools were allocated to mentors in an effort to make it easy to travel to schools for their visits. Another consideration was that it should be geographically or logistically (in terms of access to transport) feasible for mentors to get their TAs together for training.

3.2.4 Frequency of visits

While the TAs were permanently in schools from Monday to Friday, mentors conducted school visits from Monday to Thursday, arriving at the start of the school day. They would usually visit one school daily and observe the TAs during mathematics and Sepedi home-language lessons. Subsequently, those TAs would be given one-on-one feedback on the lesson the mentor had just observed. These feedback sessions could be individual or group sessions, and they could take place in the presence or absence of the teacher, depending on the nature of the feedback. Because mentors only observed

10 In the event of the teacher's absence, the TA was allowed to supervise the class and do revision work, provided another school teacher did regular visits. If the class being supervised by the TA was not visited by any teachers for more than 30 minutes, they were required to leave the classroom, inform the HOD, and inform their mentor. They would return to the classroom when another school teacher was present. Anecdotes suggest that in reality, TAs would sometimes be left alone with learners for longer than 30 minutes and would end up teaching the class. This was not formally reported to Funda Wande management at the time of intervention.

mathematics and home-language lessons, they would not necessarily do lesson observations for all the TAs during a particular school visit. However, TAs who did not receive an individual visit would still have the opportunity to check in with their mentor, ask questions about the content, and share challenges they were facing at the school or in their job more generally.

The intention was that each school should be visited by a mentor four to five times per term, depending on the number of schools allocated to the mentor. In a month, all the schools in the TA arm would have been visited by a mentor at least once. The frequency and nature of these school visits were critical for fostering strong relationships between the mentors and the TAs. However, the TAs did not get as much support as envisioned, due to interruptions by the Covid-19 pandemic and mentors being involved in training and other stakeholder engagements. In Term 1, schools were visited by a mentor on average three times. In Term 2, schools were visited by their mentor between three and five times. Term 3 was hardest hit with schools being visited by their mentor only one to three times. The TAs were also in communication with their mentors via WhatsApp and gathered for centralised training once per term (discussed below).

4 Structured pedagogy

The Funda Wande TA intervention falls within the ambit of a structured-pedagogy programme which is a coordinated and combined approach that consists of materials and support that structure teaching and learning for the teacher. These include a teacher guide, learner materials, training, and continuous support (Piper & DeStafano, n.d.). In this section, we discuss these elements.

4.1 Learner- and teacher-support materials[11]

According to the CAPS mathematics curriculum, by the end of Grade 1 learners should have a basic understanding of: numbers, operations and relationships; patterns, functions, and algebra; space and shape (geometry); measurement; and data handling (Department of Basic Education 2011). In literacy, learners enter the grade as non-readers and should progress to being emergent readers by the end of Grade 1 (Ardington et al. 2021). At the emergent-reader level, learners should have developed some phonological awareness, understood the alphabetic principle and known all single-letter sounds. This enables them to form syllables and words and read, write, and understand simple sentences. The mathematics[12] and Sepedi home-language materials have been collaboratively developed by Funda Wande, specialists across universities, and government curriculum advisors. All intervention schools received a set of teacher guides, learner activity books, big books, Vula Bula anthologies of Sepedi graded

11 For external evaluation reports on the mathematics and home-language materials please see Evans & Sorto (2021), and Mason & Snow (2021) respectively.

12 Bala Wande has taken a bilingual approach to language because of the recognition that the use of code-switching and translanguaging is prevalent in mathematics lessons across the country.

stories, a bilingual dictionary, posters, teacher- and learner-manipulatives. A more in-depth discussion about the development and use of these materials can be found in the Mtsatse (2022) and Sapire et al. (this volume) chapters of the Reading and Intervention Volumes in this series, respectively.

4.2 Training

The TA intervention mainly adopted a cascade approach to training. Literacy and mathematics specialists who were involved in the materials development trained the mentors and teachers on the content for the term. The mentors, in turn, trained the TAs with support from the specialists. The decision for mentors to train the TAs served both a logistical and relationship-building purpose. When mentors conducted the classroom observation, they would be able to revisit the concepts they trained the TAs on, and the TAs would, in turn, be confident in the mentors' ability to guide them. Training took two days for literacy and two days for mathematics per term (i.e. four full days of training per term for teachers, mentors, and TAs). This section summarises the content on which TAs were trained in terms 1 to 3 of 2021. A more detailed outline is available in Appendix 02.

4.2.1 Mathematics

The mathematics or Bala Wande (BW) programme focused on developing the learners' ability to understand and work with numbers (number sense). This necessitated emphasis on the numbers, operations and relationships domain. The programme placed a lot of weight on conceptual understanding through the whole-class learner activities and on procedural fluency through the mental-maths activities on which TAs were extensively trained. Learning was facilitated by reinforcing previous lessons and building on the prior knowledge that was needed to meaningfully take part in the activities.

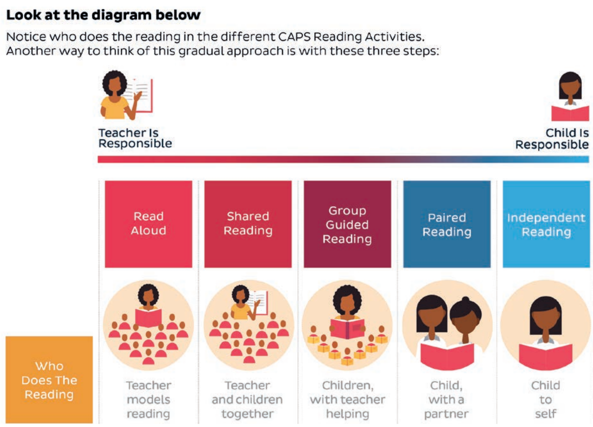
The content covered moved progressively from creating foundations for number sense in term 1, introducing addition and subtraction in term 2, to the implementation of a 'catch-up programme' in term 3. Learners were scaffolded by first being taught to recognise and write numerals, order and compare numbers, understand number bonds, work with two-dimensional shapes (recognising, counting, and sorting), and recognise and understand the reasoning behind patterns. Mathematical operations were introduced in the second term when the foundations of number sense should have been consolidated. Term 2 thus progressed from concepts of increasing and decreasing (and the relationship to addition and subtraction), to the use of a number track and the processes of counting all (starting from 1) and counting on (starting from the number they are adding to), to the use of number bonds. TAs were trained to pay special attention to helping learners develop their mathematical vocabulary and helping learners to express their thinking by talking out loud so that the teacher could help them with their conceptual knowledge.

By term 3, it was evident from internal monitoring that teachers were struggling to keep up with the curriculum. In response to this, a Read, Write, Calculate Catch-up programme was implemented, which revisited and consolidated the concepts of number bonds, addition, and subtraction that were taught in the previous term. Furthermore, both teachers and TAs were trained on strategies to help struggling learners, such as reflective guidance and scaffolding levels.

4.2.2 Literacy

Underpinning the literacy programme (Funda Wande, FW) are the five pillars of literacy acquisition namely, phonemic awareness, vocabulary, phonics, fluency, and comprehension (National Reading Panel, 2020). The programme followed an explicit and systematic approach to teaching phonics to emergent readers. The activities in this programme are based on a gradual-release model that moves the responsibility of reading from teachers to learners in an 'I do, we do, you do' manner over the course of five reading activities (see Figure 1). A more detailed outline of the reading activities (and the role of the TA therein) is available in Appendix 01.

Figure 1 Image of the Gradual Release of Responsibility model



Note: *Funda Wande, Reading Academy Combo Book, page 88, CC BY 4.0. From https://fundawande.org/learning-resources?i=11#learningResourcesHolder*

TAs were trained on the five reading activities in Figure 1 to build an understanding of the role of the teacher and learners in each activity, the expected outcomes of the activities and the best way to support them. TAs were expected to lead the read-aloud activity with learners but play more of a supportive role to learners in the remaining four activities.

Read-aloud: 'I do'

TAs were expected to conduct read-aloud activities during the "listening and speaking" lesson at the beginning of the Sepedi home-language period. During this activity, the TA was encouraged to read fluently, with expression and intonation. The TA was also trained to ask a set of comprehension questions while learners answered orally. In this activity, the TAs were taught how to expose learners to vocabulary that would be taught throughout the day.

Phonics and handwriting

The TAs were taught the phonics lessons that were covered in the term. Phonics was taught in a planned sequence and pace where the teacher or TA first modelled the letter-sound relationship, practised with the learners, and then allowed the individual learners to practise orally. It was important for the learner to be able to identify the letter sounds quickly and accurately regardless of the order of representation. Learners were taught to blend letters into syllables, syllables into words, and words into sentences. The phonics lesson was reinforced with handwriting or building words with word cards. TAs were also explicitly taught how to help learners with the foundations of handwriting, such as pencil grip and drawing different letter formations within writing lines.

Read, write, calculate catch-up programme

As a response to the organisational call to create a month-long catch-up programme, the literacy team focused on three components of reading: (1) decoding, (2) comprehension, and (3) response.

Decoding involves recognising phonemes (the smallest unit of sound in a language), developing letter-sound knowledge, and using this knowledge to read or sound out letters, syllables, words, sentences, etc. This should be done quickly, accurately, and with the correct intonation. The caution that was repeatedly emphasised here was that decoding is not an end in and of itself. It is a vehicle for making meaning from text (comprehension).

Comprehension involves knowing the meaning of words and understanding what they mean in the context of a sentence, paragraph, etc. Continuous scaffolding is created by linking new knowledge to existing knowledge, which opens learners up to being responsive.

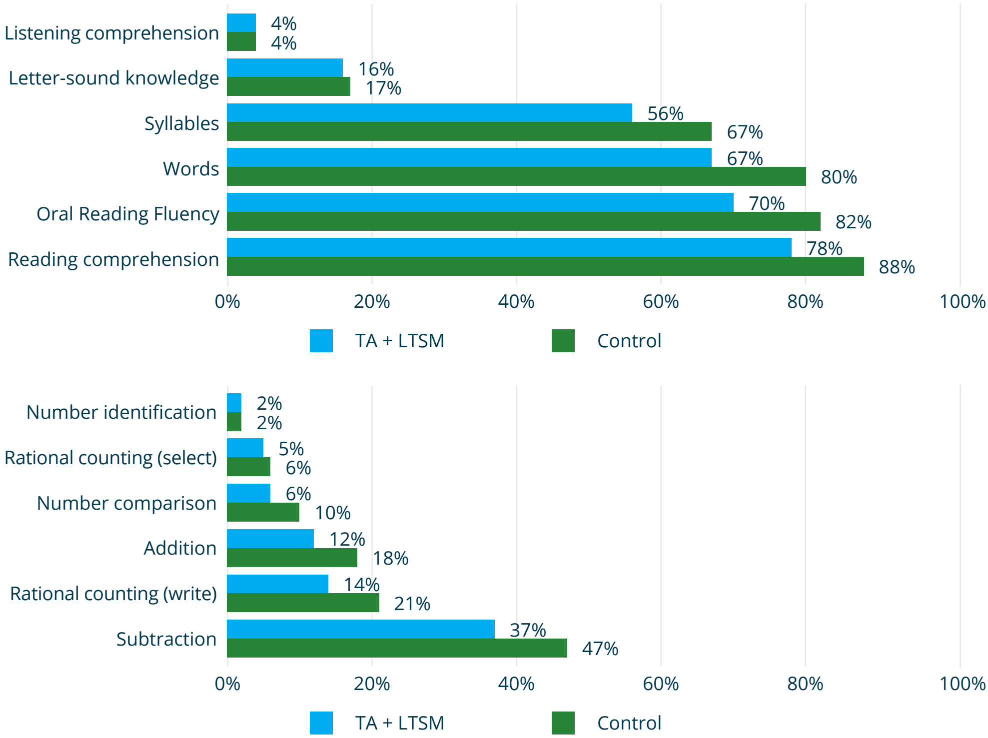
Response involves the learner producing a reaction/feeling, an oral response such as a question, or a written response such as 'What I thought about this story'. This facilitates thinking about what has been read and developing considered responses and motivates the learner to read, and allows them to gain insight and knowledge from the text they have read.

To increase the focus on the above components, the programme prioritised listening and speaking, reading, and phonics lessons to help teachers, TAs, and learners who had not been able to keep pace with the curriculum.

5 Midline results reflection

The intervention's midline evaluation ran from Term 3 Week 4 to Term 4 Week 2. Evaluators (Ardington & Henry, 2021) found that learners who received the TA intervention gained 0.44 standard deviations in their EGRA scores and 0.38 standard deviations in their EGMA scores compared to control schools. Given that the intervention was conducted during a pandemic, these results – the largest impact sizes from an RCT in education in South Africa – are especially noteworthy. In this section, we briefly discuss the results from the Funda Wande Limpopo Midline report.

Figure 2 Graphs of the percentage of learners scoring zero in EGRA (first) and EGMA (second) subtasks

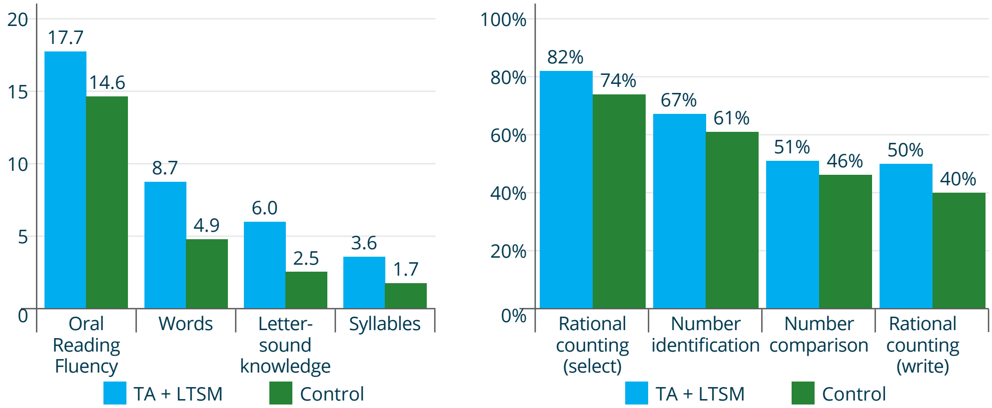


Note: *Adapted from Ardington and Henry (2021)*

A closer look at the EGRA subtasks reveals the widespread impact of Covid-19-associated learning losses (see also Ardington et al. 2021). Although intervention learners had developed their oral language skills somewhat, with the average score for listening comprehension being 55% (TA intervention) compared to 49% (control), the majority of learners scored zero on sub-tasks that assessed their decoding skills. For example, 80% of learners in control schools could not read a single word in Term 3 of Grade 1, statistically significantly greater than 67% in TA intervention schools. Pre-pandemic studies show that the percentage of learners in South Africa who could not read a single word by the end of Grade 1 was between 30% and 55% (Ardington et al., 2021). Therefore, the TAs managed to keep learners afloat amid devastating learning losses.

Similarly, in mathematics, learners were unable to achieve Grade 1-level proficiency. Learners were able to correctly engage the foundational skills that were crafted for the first term, such as the rational counting, number identification, and number comparison sub-tasks. However, they were unable to move from these foundational skills toward number sense. On average, learners in TA intervention schools were only able to correctly complete 2.4 addition number sentences of the type (2 + 3 = □) and only 1.6 subtraction number sentences of the type (3 − 2 = □) per minute. These results were statistically significantly higher than those of learners in control schools. We posit that this indicates how learners do not fully understand number bonds or are unable to translate working with number bonds to working with number sentences.

Figure 3 **Graphs of the fluency in EGRA (on left: items per minute) and percentage correct in EGMA (on right) subtasks.**



Note: *Adapted from Ardington and Henry (2021).*

Despite achieving much higher comparative results, learners in TA intervention schools are nowhere near reaching grade-level proficiency. The Read, Write, and Calculate Catch-up programme was only implemented in Term 3 – teachers and TAs had insufficient time for implementation. However, TAs have ameliorated the learning losses felt in LTSM intervention and control schools. More research is needed to understand better how TAs can effectively play a part in a learning recovery programme.

6 Recommendations

6.1 Rigorous selective recruitment

We cannot overstate the value of putting effort into recruiting the right youth to join the programme. We found that the best way to select from a large pool of job seekers was to filter through a basic set of requirements, assess whether the candidates had the desired knowledge, competencies and skills, and evaluate their interpersonal and soft skills. Initial filtering using easy-to-mark literacy and numeracy assessments allowed us to reduce the total to manageable levels before filtering on merit-based criteria. As a result of this rigorous recruitment (and a symptom of the high rate of youth unemployment in the country), we attracted many candidates who exceeded the minimum requirements. A third of the 2021 cohort of Funda Wande TAs had or were working towards post-Grade 12 qualifications, including postgraduate qualifications in education. We believe the main factors that determined the success of TAs in schools were trainability, performance, and retention.

6.2 Adaptive and responsive training

Once satisfied with the calibre of the recruited TAs, we needed to upskill them on mathematics and literacy content and classroom readiness as well as familiarise them with the programme's LTSM they would be using. This training and development of TAs had to happen in unprecedented circumstances. The team had to be especially agile and constantly adapt to the ever-changing environment with the aim of improving learner outcomes.

6.3 Ongoing monitoring and support

The dedicated team of mentors (and Funda Wande subject specialists) were invaluable. Apart from training TAs and keeping them accountable for their work, the mentors provided important psychosocial support to TAs. Mentors motivated and encouraged their TAs who were working in difficult contexts and facing the reality that many of their learners would not reach grade-level proficiency by the end of the year. Mentors also helped to mediate if there were conflicts or misunderstandings between TAs and school staff. More important than content knowledge, the key factors in building an effective support system were that the mentors conducted relatively frequent school visits (at least once a month) and had a high capacity for relationship-building.

6.4 Opportunities for growth

6.4.1 School and teacher preparedness

Schools readily welcomed TAs despite being under-prepared for a structured TA intervention. As a result, TAs were not always deployed in the most effective functions. Teachers and school management were not given introductory training focused on making use of TAs in their classrooms. Consequently, TAs were often given tasks, which did not translate into improving learner outcomes[13]. This was adapted for the 2022 intervention.

6.4.2 Limited time for revision, play and small group work

The school closures and rotational timetables put teachers under immense pressure as they tried to keep pace with the curriculum. TAs were trained on activities and games designed to scaffold and spark the learners' love for learning. These practices were often not prioritised in the classroom, even though teachers and TAs were trained explicitly on them. From internal monitoring, we found that no mathematics games were played with learners in 38% of the lessons observed by mentors in terms 3 and 4. Similarly, there was no small-group work done by the TA in 34% of the literacy lessons observed. This is concerning as it indicates that struggling learners were still

13 Conversely, in some cases, TAs conducted reading activities and other parts of the lesson that should have been reserved for the teacher. This was due to teachers leaving the classroom for more than 30 minutes at a time. This was not formally reported and is implied from anecdotal evidence.

not getting enough focused attention to learn even with the addition of a TA to the classroom.

These are some of the key pragmatic learnings we had from our first year of working in the TA schools. The intervention is currently set to end in December 2022. The return of full (and overcrowded) classrooms presents new challenges of improving learner outcomes with the presence of effective TAs in the classroom.

6.5 Wider implications: Presidential Youth Employment Initiative – Basic Education Employment Initiative

Towards the end of 2020, the government rolled out the Presidential Youth Employment Initiative – Basic Education Employment Initiative (PYEI-BEEI) as a response to high youth unemployment and the dire socio-economic conditions exacerbated by the pandemic.

> *We have committed R100 billion over the next three years to create jobs through public and social employment as the labour market recovers. [...] We are going to create 300,000 opportunities for young people to be engaged as education and school assistants at schools throughout the country*
>
> President Cyril Ramaphosa (October 2020).

The Department of Basic Education has been the biggest beneficiary of this economic stimulus package. Although it was initially meant to be an employment initiative, there are real opportunities to improve learning outcomes. Given that the PYEI-BEEI initiative was rolled out in 2021, these 'educator assistants' (who are different from the TAs discussed in this chapter) were also in treatment- and control-schools. While this programme holds much promise for large-scale impact on learning outcomes, it is our view that that promise depends on more rigorous selection, training, and support in line with the programme documented in this chapter and the provision of a structured pedagogy programme with aligned materials. Through continuous experimentation at a small scale and engagements with the government, there is an opportunity to show how youth can be employed in a nationwide learning-recovery programme.

7 Conclusion

The Coronavirus pandemic exacerbated the already severe gap between where learners should be and where most learners are at the end of each grade. The Funda Wande Limpopo TA intervention was implemented while learners around the country were experiencing severe learning losses. At the end of one year of the intervention, learners in TA intervention schools made substantial gains in both mathematics and home-language compared to schools that continued with business as usual (control schools). Some of the supporting factors behind this were the recruitment of highly capable and well-suited TAs, strong relationships between TAs and their mentors, close monitoring and support, and high-quality materials and training developed by Funda Wande content specialists.

Even after the significant gains in these intervention schools, the picture remains a sobering one, as learners were still far from reaching grade-level proficiency at the time of evaluation in Terms 3 and 4. There is space for improvement in terms of programme fidelity. Learners are still not getting enough focused attention and opportunities for remediation through small group activities and games. Furthermore, it is yet to be determined if the end of rotational timetabling will harm learners (due to larger class sizes) or help them (through better curriculum coverage). Nevertheless, the lessons learned from the design and implementation of the TA intervention are encouraging and can be applied more broadly to the education system.

References

Ardington, C., & Henry, J. (2021, December). *Funda Wande Limpopo Evaluation* [Midline Report]. South Africa Labour and Development Research Unit (SALDRU). https://fundawande.org/img/cms/news/Limpopo%20TA%20Evaluation%202021.pdf (Accessed 14 November 2022).

Ardington, C., Wills, G., & Kotze, J. (2021). COVID-19 learning losses: Early grade reading in South Africa. *International Journal of Educational Development*, 86(102480).

Ardington, C., Wills, G., Pretorius, E., Mohohlwane, N., & Menendez, A. (2021, July). Benchmarking oral reading fluency in the early grades in Nguni languages. *International Journal of Educational Development*, 84(102433).

Banerjee, A., Banerji, R., Berry, J., Duflo, E., Kannan, H., Mukherji, S., Shotland, M., & Walton, M. (2016). *Mainstreaming an Effective Intervention: Evidence from Randomized Evaluations of "Teaching at the Right Level" in India* (Issue w22746). National Bureau of Economic Research.

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

Department of Basic Education. (2011). *Curriculum and Assessment Policy Statement: Grades 1-3, Mathematics*. Pretoria: Department of Basic Education

Duflo, A., Kiessel, J., & Lucas, A. (2020). *Experimental Evidence on Alternative Policies to Increase Learning at Scale* (Issue w27298). National Bureau of Economic Research.

Evans, N., & Sorto, A. M. (2021). *Bala Wande Grade 1 Teacher Guide and Learner Activity Book* [Desk Review]. https://fundawande.org/img/cms/news/Funda%20Wande_Maths%20Report.pdf (Accessed 14 November 2022).

Howie, S. J., Combrinck, C., Roux, K., Tshele, M., Mokoena, G., & McLeod, P. N. (2017). *PIRLS Literacy 2016: South African Highlights Report (Grade 4)*. Centre for Evaluation and Assessment.

Mason, P. A., & Snow, C. E. (2021). *Funda Wande Language and Literacy Program* [Research Report]. Available at: https://fundawande.org/img/cms/news/Harvard%20Report_1.pdf (Accessed 14 November 2022).

Mtsatse, N. (2022). Developing the Funda Wande Literacy and Life Skills Grade 1 Workbook and Teacher Guide. In Spaull, N. & Pretorius, E. (Eds) *Early Grade Reading in South Africa*. Cape Town: Oxford University Press.

National Reading Panel. (2020). *Teaching children to read: an evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups*. National Institute of Child Health and Human Development, National Institutes of Health. Available at: https://www.nichd.nih.gov/sites/default/files/publications/pubs/nrp/Documents/report.pdf (Accessed 14 November 2022).

OECD. (2020). TALIS 2018 Results (Volume II): T*eachers and School Leaders as Valued Professionals*. Paris: OECD Publishing.

Piper, B., & DeStafano, J. (n.d.). Structured Pedagogy Series. Available at: https://scienceofteaching.site/how-to-guides/learning-outcomes/topic/lesson-1/ (Accessed 14 November 2022).

Ramaphosa, C. (2019). The Presidency, Republic of South Africa. *State of the Nation Address*. Parliament. https://www.stateofthenation.gov.za/assets/downloads/SONA-June-2019.pdf (Accessed 14 November 2022).

Ramaphosa, C. (2020). *South Africa's Economic Reconstruction and Recovery Plan* [Address by President Cyril Ramaphosa to the Joint Sitting of Parliament on South Africa's Economic Reconstruction and Recovery Plan]. Available at: https://www.gov.za/speeches/president-cyril-ramaphosa-south-africa%E2%80%99s-economic-reconstruction-and-recovery-plan-15-oct (Accessed 14 November 2022).

Reddy, V., Winnaar, L., Arends, F., Harvey, J., Hannan, S., Namome, C., & Zulu, N. (2020). *TIMSS 2019 Highlights of South African Grade 5 Results in Mathematics and Science*. Department of Basic Education. Available at: https://www.timss-sa.org/publication/timss-2019-highlights-of-south-african-grade-5-results-in-mathematics-and-science (Accessed 14 November 2022).

Shepherd, D., & Mohohlwane, N. (2021). A generational catastrophe: COVID-19 and children's access to education and food in South Africa. *Development Southern Africa*. 39(5), 762–780.

Spaull, N., Pretorius, E., & Mohohlwane, N. (2020). Investigating the comprehension iceberg: Developing empirical benchmarks for early grade reading in agglutinating African languages. *South African Journal of Childhood Education, 10*(1), 1–14

Spaull, N., Courtney, P., Qvist, J., (2022). Mathematical stunting in South Africa: An analysis of Grade 5 mathematics outcomes in TIMSS 2015 and 2019. In Venkat, H., & Roberts, N. (Eds) *Early Grade Mathematics Research & Interventions in South Africa*. Cape Town: Oxford University Press.

Van Staden, S. & Gustafsson, M. (2022). What a decade of PIRLS results reveals about early grade reading in South Africa: 2006, 2011, 2016. In Spaull, N. & Pretorius, E. (Eds) *Early Grade Reading in South Africa*. Cape Town: Oxford University Press.

appendix /01

Recruitment of Teaching Assistants

Funda Wande
Reading for Meaning

We're Hiring!

Teacher Assistant Mentors

LOCATION: POLOKWANE

Overview

Funda Wande is a dynamic not-for-profit organisation that aims to ensure that all learners in South Africa can read for meaning (Literacy) and calculate with confidence (Maths) in their home language by the age of 10. Reporting to the Project Manager, the Teacher Assistant Mentor (TAM) will help train and guide TAs on early grade learning content. They will also provide psychosocial support to TAs and help them navigate challenges working with learners and schools.

Job responsibilities and role:

Mentoring

- Develop TAs' skills and behaviours through professional development provided by Funda Wande. This includes facilitating and coordinating training and workshops for TAs.
- Provide technical support to TAs on the Funda Wande website and APP. Help TAs with the use of Funda Wande resources and materials to support teachers' instruction and planning.
- Conduct lesson observations and provide feedback to TAs on how to better assist teachers and support learners.
- Develop and maintain relationships with teachers, school management teams, subject advisors, and other TAMs. Attend workshops related to Funda Wande work.

Administration, Planning & Reporting

- Prepare weekly/monthly work schedules and share them with the Project Manager.
- Submit data, logs and documentation related to mentoring.
- Prepare weekly reports detailing activities completed, challenges faced, and achievements reached by TAs.
- Thoroughly plan and facilitate meetings TA feedback meetings.

Experience and Qualifications:

- Degree in Education (Foundation Phase).
- Minimum 2-year experience teaching in Foundation Phase or working with Foundation Phase aged children.
- Proficient in Microsoft Office, Google Suite, and related applications.
- Proficient in Sepedi and English (both oral and written) and can communicate Foundation Phase Maths in Sepedi.
- A valid driver's license and own reliable transport.
- Experience working with youth is essential.

Behavioural competencies:

- Models continual improvement, demonstrates lifelong learning, and applies new learning.
- Demonstrates leadership qualities and strong interpersonal skills, including giving and receiving constructive feedback. Ability to work effectively as an individual and as a member of a team. Shows ability to take initiative.
- Ability to adapt to changing organisational needs, conditions, and work responsibilities as required.
- Strong conflict management/mediation skills.

Contract:

18 months with the possibility of extension.

Applications:

To apply for this position, please submit the following to applications@fundawande.org

- A 1-page motivation letter
- A 3-page (maximum) CV

Deadline for applications:

29 January 2021

If you have not heard back from us by the 28 February 2021, please consider your application unsuccessful.

appendix /02

Content of Teaching Assistant Training

The following section expands on section 4.2 of this chapter, giving a more detailed description of the content on which TAs were trained.

Mathematics

Creating foundations for Number Sense

The focus of Term 1 was largely on the number subdomain. The content was structured to build the foundations of number sense in the 1 to 10 number range. Learners were scaffolded by first being taught to recognise and write numerals, ordering and comparing numbers, understanding number bonds, working with two dimensional shapes (recognising, counting, and sorting), and being able to recognise and understand the reasoning behind patterns. Mathematical operations were introduced in the second term when the foundations of the numbers should have been consolidated.

Introducing addition and subtraction

The beginning of Term 2 focused on the concepts of increasing and decreasing as well as the relationship to addition and subtraction. Teachers and TAs were introduced to the number track where they were explicitly trained on how the activity helps the learner from counting all (starting from 1), to counting on (starting from the number they are adding to), which would progress to using number bonds. TAs were trained to pay special attention to helping learners develop their mathematical vocabulary and helping learners to express their thinking by talking out loud so that the teacher would help them with their conceptual knowledge.

Addition number sentences (e.g. $2 + 3 = \square$) were introduced in the second week of Term 2. The concepts that were developed involved join type problems (combining two quantities to make a total), compare type problems (similarities and differences between quantities), and interpreting oral word problems. Similarly, subtraction number sentences were introduced in week 4. Here teachers and TAs were trained on change type problems (subtracting a smaller quantity from a larger quantity) and part-whole type problems (separate the whole into two parts). These concepts were meant to be consolidated within the first six weeks of the term. The remaining four weeks were dedicated to introducing measurement, sorting three dimensional objects, and understanding repeated geometric patterns.

Read, write, calculate catch-up programme

By Term 3 it was clear that the rotational timetabling was making it near impossible for teachers to keep up with the curriculum. Since one of the tenets of Funda Wande is to prioritise children's learning over keeping pace with the curriculum, the organisation implemented a 'Read, Write, Calculate Catch-up programme'. This meant that the first month of Term 3 was dedicated to consolidating concepts that should have been learned in the first two terms. Teachers could resume the planned instruction for the remainder of the term after this first month of catch up with the approval of the province and district education departments.

The Read, Write, Calculate Catch-up programme revisited and consolidated the concepts of number bonds, addition, and subtraction that were taught in the previous term. Furthermore, both teachers and TAs were trained on strategies to help struggling learners such as reflective guidance and scaffolding levels according to the Concrete-Pictorial-Abstract (CPA) strategy (Jordan, Miller, & Mercer, 1998) which are discussed below.

When reflective guidance is applied correctly, learners are given the opportunity to improve their mathematical thinking. There are three components of this strategy: (1) noticing and engaging the errors that learners make, (2) asking learners questions that will encourage them to evaluate and correct their reasoning instead of giving them the answers, and (3) understanding the value of scaffolding levels tool (Figure A1) to help learners improve their reasoning and conceptual knowledge.

Figure A1 Strategy of using Concrete, Pictorial, Abstract (CPA) scaffolding levels

Source: Funda Wande, Limpopo Grade 1 Term 3 Training slides, CC BY 4.0.

Literacy

The literacy programme applied the spiralling model to scaffolding the learning progression for children. Concepts were introduced through the theme of the week and revisited several times while increasing the complexity of each revisit. This was designed to give learners multiple opportunities to be exposed and practice new knowledge while building on existing knowledge.

Shared reading: 'We do'

In this activity, the responsibility of reading shifts to mutual participation of the teacher and learners. The teacher should first model fluent reading then learners join in while the teacher points under the printed text. Comprehension and decoding questions should be asked in a manner that builds on prior knowledge. During this time, the TA was meant to assist the teacher by observing the class, helping with classroom management and identifying struggling learners.

Group guided reading: 'We do' → 'You do'

The responsibility in this activity shifts further towards the learner while the teacher guides them. Following an assessment of the learners' abilities, the teacher creates small groups of learners of similar reading ability. It is in this activity that the teacher gets a chance to give learners dedicated time to read at their level (Fountas & Pinnell 2012). Each learner in the group is given an opportunity to engage with the chosen text. During this time, the teacher should concentrate on developing the learners' phonemic awareness, vocabulary, fluency, and comprehension. The other learners in the group are then expected to follow by reading silently. While the teacher is busy working with the small group, the TA should serve the dual purpose of managing the rest of the class and helping learners with their independent work.

Independent work and paired reading: 'You do'

These activities took place for the rest of the class during group guided reading. During independent work, learners completed a set of tasks in their Learner Activity Books that allowed them to practise letter writing and recognition, phonics, and vocabulary. In the third term, paired reading was introduced for learners who were able to carry out this activity.

Table A1 Intervention training schedule for Terms 1 to 3

<table>
<tr><th>Duration
(medium)</th><th>Attendees
(percentage)</th><th>Facilitators</th><th>Content</th></tr>
<tr><td colspan="4">TERM 1</td></tr>
<tr><td colspan="4">Literacy</td></tr>
<tr><td>2 days
(in-person)</td><td>5 TA Mentors
(83%)</td><td>2 Literacy
Specialists</td><td rowspan="3">• Introduction to Funda Wande:
• The Workbook Approach
• Integration of Home Language & Life Skills
• Key Practices for Read Aloud.</td></tr>
<tr><td>2 days
(in-person in 6 hubs)</td><td>76 TAs (99%)</td><td>5 TA Mentors</td></tr>
<tr><td>2 days
(in-person in 2 hubs)</td><td>46 Teachers
(60%)</td><td>2 Literacy
Specialists</td></tr>
</table>

Duration (medium)	Attendees (percentage)	Facilitators	Content
Maths			
2 days (virtual via Zoom)	5 TA Mentors (83%)	1 Maths Specialist	• Introduction to Bala Wande • Mental Maths activities • Games • Concept Development • Learner Activity Book activities.
2 days (in-person in 6 hubs)	76 TAs (99%)	5 TA Mentors	
2 days (in-person in 2 hubs)	46 Teachers (60%)	2 Maths Specialists	
TERM 2			
Literacy			
2 days (in-person)	6 TA Mentors (100%)	2 Literacy Specialists	• Integration of Home Language & Life Skills • Read Aloud demonstrations • Weekly overview of Big Book, Learner Activity Book, and Teacher Guide
2 days (in-person in 6 hubs)	73 TAs (95%)	6 TA Mentors	
2 days (in-person in 2 hubs)	72 Teachers (94%)	3 Literacy Specialists	
Maths			
2 days (virtual via Zoom)	6 TA Mentors (100%)	1 Maths Specialist	• Addition (increase and decrease) • addition stories and patterns • Introducing Subtraction
2 days (in-person in 6 hubs)	73 TAs (95%)	6 TA Mentors	
2 days (in-person in 2 hubs)	72 Teachers (94%)	3 Maths Specialists	
TERM 3			
Literacy			
2 days (in-person)	6 TA Mentors (100%)	2 Literacy Specialists	• SMART Approach to LTSM, and Pedagogy • Assessment & Training • Principles of reading (Decoding, Comprehension & Response) • Explicit and systematic teaching of Phonics • Group Guided Reading and Assessment • Read, Write and Calculate Catch-up Program
2 days (virtual via WhatsApp)	77 TAs (100%)	6 TA Mentors	
2 days (in-person in 2 hubs)	55 Teachers (71%)	3 Literacy Specialists	
Maths			
2 days (virtual via Zoom)	6 TA Mentors (100%)	1 Maths Specialist	• Addition (increase and decrease), • addition stories and patterns, • Introducing Subtraction, • Subtraction Problems and Patterns, • Subtraction Stories and Patterns
2 days (virtual via WhatsApp)	77 TAs (100%)	6 TA Mentors	
2 days (in-person in 2 hubs)	55 Teachers (71%)	3 Maths Specialists	

09

Using technology to improve English literacy: The case of ‘Reading Eggs’ in South Africa 2012-2021

JOY OLIVIER, NICOLA HARRIS, MEGAN BOROLE & BRUCE MCDOUGALL

Abstract

Educational technology (ed-tech) holds potential to deliver and assess learning at scale, and yet is notoriously difficult to implement, with most interventions delivering disappointingly low levels of impact, if any. This chapter details key aspects of Click Learning’s English literacy intervention, leveraging ed-tech in under-resourced South African schools, and highlights the lessons learned to date. In partnership with over 224 schools, Click has established and runs computer laboratories in Gauteng, the Eastern Cape, KwaZulu-Natal, Mpumalanga and the Western Cape. Click deploys the Reading Eggs suite of English-literacy online applications to more than 125,000 learners in Grades R-7. During scheduled lessons within the school timetable, learners work through the applications’ activities at their own pace. In this chapter we document the number of hours learners use the applications, and present data for the years 2018 to 2021. Data on the intervention’s implementation to date is followed by a discussion of strategies for impact evaluation going forward. Click has developed an eQuiz literacy assessment, which holds promise for low-cost learner assessment at scale. By explaining how the programme works in under-resourced schools and documenting programme usage and learner outcomes, we hope to contribute to the growing literature on the potential for ed-tech interventions to improve literacy outcomes, and to inform strategies for learning loss recovery in the wake of the pandemic.

KEYWORDS
educational technology (ed-tech), English as First Additional Language (EFAL) foundation phase, early grade reading, English literacy, catch-up programmes

1 Introduction

Educational technology (ed-tech) holds potential to deliver and assess learning at scale, and yet is notoriously difficult to implement, with most interventions delivering disappointingly low levels of impact, if any. Despite the rapid growth of internet access and technology use worldwide, including in low-income and middle-income countries (United Nations 2019), as well as the proliferation of e-learning applications (Grand View Research 2022). While ed-tech has been deployed in developing countries, little has led to any measurable improvements in learner outcomes.

This chapter begins with a review of the international and local literature, focusing on available evidence pertaining to the provision of Information and Communications Technology (ICT) in schools, and the impact of ed-tech literacy interventions. The review explores the impact of ed-tech on early grade literacy outcomes, with a focus on initiatives that involve structured ed-tech interventions employing self-led software applications used directly by learners.

Next, in describing Click Learning's (Click's) approach, the chapter details key aspects of the intervention, and highlights the lessons learned in leveraging ed-tech in under-resourced South African schools to date. Click deploys the Reading Eggs suite of English literacy online applications to learners in Grades R-7 in under-resourced schools. During scheduled lessons within the English as First Additional Language (EFAL) and English as Home Language (HL) school timetable, learners work through the applications' activities at their own pace. Click has established and runs computer laboratories in 224 schools in Gauteng, the Eastern Cape, KwaZulu-Natal, Mpumalanga and the Western Cape, and delivers the programme to more than 125,000 active learners.

In this chapter we highlight enabling factors for implementing ed-tech in South Africa's no- and low-fee schools, and outline future possibilities for measuring learners' literacy outcomes at scale.

2 Literature

2.1 International literature

Learners who are unable to read well in early grades are at risk of poor performance in other subjects and grades, emotional and behavioural problems, and school drop-out (Lesnick, et al. 2010). All over the world, interventions seeking to improve foundational literacy and numeracy have employed ed-tech with mixed results. In this paper, the authors adopt Cheung and Slavin's (2013, 279) working definition of ed-tech as "a variety of electronic tools and applications that help deliver learning content and support the learning process".

Evidence of the impact of ed-tech on learning outcomes is inconclusive (Skryabin et al. 2015), and success is often contingent on how and under what conditions interventions are implemented (Ndlovu et al. 2021). In their study drawing on large-scale learner data – from the Trends in International Mathematics and Science Study

(TIMSS) 2011, Progress in International Reading Literacy Study (PIRLS), 2011 and the Programme for International Student Assessment (PISA), 2012 – Skryabin et al. (2015, 57) investigated the influence of individual ICT usage on learner achievement, with inconclusive and complex findings showing mixed influences across subjects and grades, "depending on the different types of ICT usage".

Rodriguez-Segura (2021) shows that in some cases learning gains from ed-tech programmes delivered after traditional classroom hours were due to the additional learning time, rather than the ed-tech. He points to the case by Ma et al. (2020, 2), which isolates and tests the technology-based effects of computer-assisted learning, through a novel multi-treatment field experiment with more than 4,000 children in rural China. While the researchers found evidence of positive overall programme effects on academic outcomes, when the tech-based effects of computer-assisted learning were isolated, small to no effects were found. The authors note that often computer-assisted learning programmes are held after school, and include non-technology-based inputs such as facilitators, and conclude that "at times, the 'tech' in ed-tech may have relatively small effects on academic outcomes" (Ma et al. 2020, 2).

While the numbers of ed-tech interventions seem to proliferate exponentially, unfortunately very few are evaluated. For example, of the 1,640 featured in Brookings' Centre for Universal Education's global catalogue, just 11% had been externally evaluated, and only 26 (1.6%) with randomised control trials (RCTs) (Vegas et al. 2019). In a rigorous review of 26 studies on the use of hardware in teaching, Muralidharan, Singh and Ganimian (2018) find that the bulk of studies (67%) in 11 countries find no result at all (40%) or an ambiguous result with some positive and some negative results (27%). Only 27% of the studies found positive results and a minority found negative results (6%). Sample sizes for studies with 'no impact' were often large and programme intensity was substantial (Muralidharan et al. 2018, Appendix C).

2.1.1 Ed-tech for self-led learning

The type of interventions described as ed-tech are very wide-ranging: from WhatsApp messaging parents or teachers, to online teacher development, and structured self-led learning (where Click's intervention sits). In his paper on the evidence of ed-tech in developing countries, Rodriguez-Segura (2021) writes that structured, self-led learning interventions are among the most effective at raising learning outcomes. Muralidharan et al. (2018, 1) identify "the mismatch between the level of classroom instruction and student learning levels" as a key likely cause of developing countries' failure to transform universal access to primary education into universal literacy and numeracy for primary school learners. While there is a proliferation of convincing evidence on the success of the teaching-at-the-right-level methodology (TaRL) in addressing the heterogeneity of learning levels in classes (Vromant et al. 2021), the ways in which ed-tech can contribute in this regard are for the most part unknown.

Self-led-learning ed-tech aims to address learner-specific gaps by adjusting the level and pace of instructional content in response to learner engagement with the software. In his review of ed-tech evaluations, Rodriquez-Segura (2021) finds consistently medium to large effects for learning outcomes from interventions in developing countries employing self-led-learning software. He notes that these interventions are often cost-effective (as laboratories in schools and communities can serve multiple learners), and are best used to complement classroom instruction and fill content gaps. However, he also cautions that the reliance on self-guidance may better suit higher achievers, and hence increase within-class inequality.

In their study of Mindspark, a tech-enabled after-school programme in India, Muralidharam et al. (2018) found that learners made large gains in both maths and language, to which the researchers attribute the software's personalised learning and feedback. In this intervention, children in urban India used a personalised, self-led-learning application to which they had won free access via lottery. The centres provided 90 minutes of daily programming, six days per week. Each session comprised 45 minutes of computer-based learning that is self-led and 45 minutes of instructional support from a teaching assistant in groups of 12-15 students for 4.5 months, i.e. a total of 162 hours of engagement with the blended programme. The researchers found that the programme had positive effects and was cost-effective for both maths and Hindi across different grades. While the authors note that they could not isolate the ed-tech aspect of the initiative in their analysis, they surmise that as similar research by Berry and Mukherjee (2016) on an initiative providing after-school instruction to learners of the same age found no impact on test scores, that the programme effects are "most likely attributable to the computer-assisted learning component" (2016, 1429).

2.1.2 Ed-tech for literacy

In their meta-analysis of 84 studies with high methodological standards, and covering more than 60000 K-12 learners in the United States, Cheung and Slavin (2012) found that most classroom-based ed-tech programmes did not produce meaningful effects in reading, and the higher the methodological quality, the lower the effect size. Cheung and Slavin (2013) also reviewed 20 studies on the effectiveness of ed-tech in improving the reading achievement of struggling learners in American elementary schools, and found an overall positive, although small, effect. Piper et al. (2016) implemented an RCT to evaluate three interventions for early grade reading in Kenya (e-readers for learners, tablets for teachers, and the base Kenya Primary Math and Reading (PRIMR) programme with tablets for instructional supervisors). The researchers found that the ICT-enabled interventions did not improve literacy outcomes significantly more than the base PRIMR programme without ICT, and were less cost-effective.

Comings (2020) presents a number of case studies of ed-tech interventions aiming to improve early grade reading achievement. Included in this review is his (2018) evaluation of "Antura and the Letters" and "Feed the Monster" in Arabic, which found weak but encouraging evidence that the use of smartphone learning apps can build basic Arab literacy skills and improve the psychosocial wellbeing of Syrian children. The evaluation unfortunately suffered from high dropout rates due to implementation in a crisis context, and hence sample sizes were small, and learners only used the application for an average of 22 hours. Using the Early Grade Reading Assessment (EGRA) instrument, evaluators found statistically significant improvements in syllable reading, invented word, and oral-reading fluency. Comings (2020, 8) writes that given the small dosage (i.e. the number of application usage hours), "the percentage gains and effect sizes are encouraging and suggest that a larger dosage might produce meaningful gains".

Another case described by Comings is the E-books 4 Khmer project (E4K) in Cambodia, evaluated by School-to-School International (2017), which transformed reading textbooks into levelled e-books with embedded gamified quizzes for learners to work through at their own pace. The researchers found that students with access to the application had statistically significant greater gains on the EGRA oral-reading fluency and reading comprehension subtests. Although Comings expresses some misgivings on these findings due to a difference in pre-test scores between treatment and control groups, he notes that the large differences in gains may indicate that the intervention had a positive impact.

Comings also presents two cases disseminating local stories and comprehension questions to learners in Zambia and India. In Zambia, Creative Associates International implemented the Makhalidwe Athu ('Our Way of Staying') project (2019), which sent locally crowd-sourced stories in ciNyanja via SMS to families for 10 months. Children received a story and a related set of comprehension questions three times per week, which were also available in audio format through a free telephone call. NORAC (2018), at the University of Chicago, conducted an evaluation employing a difference-in-difference approach and found a statistically significant impact on oral reading fluency and reading comprehension subtests, with an effect size of 0.27 and 0.23, respectively. In India, Sesame Workshop India Trust (SWI) implemented the Play.Connect.Learn project (2019) in six districts in Maharashtra over 28 months. The project made stories in Marathi available, with accompanying comprehension games and quizzes. An evaluation conducted by School-to-School International found statistically significant gains in reading fluency, but not in comprehension.

In their recommendations for cost-effective learning interventions, the Global Education Evidence Advisory Panel (2020) highlight software that "adapts to the learning level of the child (where hardware is already in schools)" as a "good buy," and note that whether implemented during or after school, the approach can significantly improve learning outcomes, as evidenced by studies in Uruguay (Perera & Diego 2017) and India (Banerjee et al. 2005).

The prevailing message from studies on foundational-literacy ed-tech is that in order to impact reading outcomes, it needs to be integrated into an effective learning programme, and needs to be "well-aligned with an underlying pedagogic intervention" (Piper et al. 2016). Some authors also note the importance of analysing the cost-effectiveness of these interventions, and of comparing their effects with less expensive programmes that do not include ICT investments.

2.2 South African literature

2.2.1 English literacy in South Africa

Many learners in South Africa struggle with the shift from mother-tongue instruction to English as the Language of Learning and Teaching (LoLT) in Grade 4 (Pretorius & Spaull 2016). In 2016, 23% of Grade 4 learners who participated in PIRLS were assessed in English (the largest proportion of participating learners, followed by 21.8% who were assessed in Zulu) (Howie et al. 2017). While this group was the highest-performing out of all test languages, with a mean achievement score of 372 points, more than half of these learners were unable to attain the study's lowest benchmark. Only 21% of the Grade 4 learners who were assessed in English spoke it at home, and there was a significant difference in mean achievement scores between these learners (who had a mean score of 445) and those who speak another home language (for whom the mean score was 356).

2.2.2 Ed-tech in South Africa

In their report on "Learning with technology in low-income households in times of disruption", Ndlovu et al. (2021, 3) note the minimal impact of Covid-19 on well-resourced South African schools, and attribute this in part to differentiated ICT access. The authors propose that "while there are many hurdles to overcome in getting ICT into poor communities, it must become a weapon in our arsenal to address learning inequity".

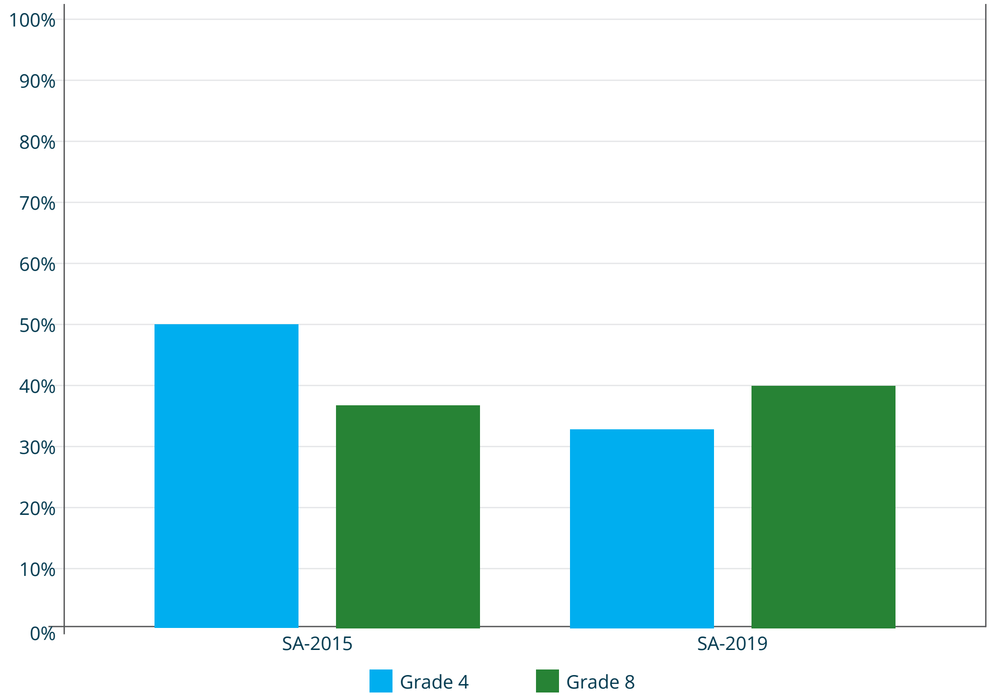
While the provision of ICT in schools is in most cases insufficient for achieving learning gains, reliably working and connected computer laboratories are an important first step, which in many developing countries is met with multiple obstacles. Karsenti et al. (2012, 2) identify the following types of challenges in implementing ed-tech in 12 African countries (including South Africa): infrastructural (e.g. "power outages and internet blackouts"); technological (e.g. "insufficient and outdated computer equipment"); human (e.g. "lack of techno-pedagogical skills and training"); and financial (e.g. "lack of permanent funding for pedagogical ICT integration").

In a Portfolio Committee presentation to Parliament in March 2020, the Department of Basic Education (DBE) reported that 71% of public schools had internet connectivity, and that 92% of these were low-speed connections. The same presentation reports the percentage of connected schools in 2018/2019 at 69% (DBE 2020). However, in the 2019 National Education Infrastructure Management System (NEIMS) report, the DBE (DBE 2019) reported that 4,695 schools (20%) have internet connectivity for teaching and learning, while 6,770 (29%) are connected for administrative purposes. They also note in the same report that 36% of ordinary operational schools have a computer centre (the report unfortunately does not provide information on the ratio of computers to learners).

Perhaps the most reliable information about learners' access to computers at school is the recent 2019 Trends in International Mathematics and Science Study (TIMSS), which surveyed nationally representative samples of Grade 5 (Reddy et al. 2020) and Grade 9 learners (Reddy et al. 2020a). Figure 1 presents the study's findings on the percentage of learners in schools with access to computers, and shows that Grade 4 computer access has dropped from 50% in 2015 to 37% in 2019, while Grade 8 access has improved marginally from a very low 33% in 2015 to 40% in 2019.

As further evidence of the difference between de jure and de facto access to computers, the ePIRLS exercise in South Africa in 2015/16 is instructive. Combrinck and Mtstase (2018) document the sampling process for ePIRLS (PIRLS on a computer) in Gauteng. The province was selected to be surveyed given its wealth and high number of English-LOLT schools in the country (ePIRLS was only done in English).

Figure 1 South African learners' access to computers at school



Source: Own calculations using TIMSS-N 2015 and 2019

The selection process began with the identification of English-LoLT primary schools with workable computer laboratories. The researchers found that although the original list provided by the Gauteng Department of Education (GDE) contained 2,161 schools, most of these were high schools or adult centres. After eliminating these, only 236 primary schools (allegedly) equipped with computer laboratories remained. After contacting the principals to verify whether they did in fact have working laboratories, the list was further reduced to just 36 schools. Of these, 25 were sampled and visited and in reality only 15 had working laboratories (Combrinck & Mtstase 2019, S4). Of Gauteng's approximately 1,600 primary schools, nearly half are English-LoLT[1]. It is striking that of nearly 800 English LOLT primary schools in Gauteng (arguably the country's wealthiest province), only 36 (5%) had functional computer laboratories.

1 Of the 4604 Grade 3 learners assessed in Verification Annual National Assessment (V-ANA) in 2013, 48% reported an English LOLT in Gauteng (Nic Spaull, Personal Communication, 10 April 2022).

Given the multiple challenges, the ePIRLS study did not meet the sampling requirements to be featured in the International ePIRLS report, and is instead viewed as an exploratory case study (Howie et al. 2017b). It is interesting to note that in schools where the computer laboratory was used, both paper-based and online-reading scores were significantly higher (93-100 points) than in schools where this was not the case. The authors surmise that "in addition to formal teaching and possible effective use of the computer room, the use of the computer room may be an indicator of the overall functionality of the school, good management, and availability of many other resources" (Howie et al. 2017b, 3).

Ed-tech research in South Africa is limited, and in most cases the findings are far from encouraging (Chigona et al. 2010; Smith & Hardman 2014). However, in their RCT of Numeric's after-school ed-tech programme, Böhmer et al. (2014) found large gains on basic numeracy outcomes, as well as higher performance in Grade 8 curriculum questions at endline; pointing to the potential for learner-responsive ed-tech to fill content gaps that classroom instruction may not. It is also worth noting, given the focus of this chapter, that while the study found no significant differential treatment effect by gender, race, home language, baseline typing speed, or cognitive development, treatment learners with better English literacy at baseline scored significantly higher than control learners in the bottom third on core Grade 9 curriculum maths questions.

Castillo (2017) conducted an RCT to investigate the impact of a computer-based early grade reading intervention for improving literacy outcomes in South Africa, and to explore whether there is a correlation in the transfer of early-literacy skills to a first additional language. The intervention delivers the Bridges to the Future Initiative (BFI) computer-based Grade 1-3 curriculum in Sepedi, Tshivenda, Xitsonga, and English. The application features self-paced, interactive audio-visual multimedia that enable the learner to hear, identify, and manipulate phonemes; play grammar, word-building and sentence-building games; and answer comprehension questions. The software also provides continuous assessment and feedback, and BFI teachers provide writing and speaking exercises guided by lesson plans. The results demonstrate convincingly that learners in intervention schools performed significantly better on mother-tongue reading fluency measures, as well as in comprehension, and indicate that employing guided and contextualised digital material to teach literacy can lead to early-reading skills development. Castillo's research emphasises the importance of establishing baseline reading skills in a mother-tongue language for improving transfer of literacy skills to English, although he found a ceiling in the extent to which mother-tongue reading skills translate into English literacy skills once higher levels of mother-tongue decoding proficiency are achieved.

There is a dearth of literature on the use of ed-tech for improving early grade literacy in South Africa. The country's abysmal pre-Covid-19 literacy outcomes have been compounded by learning loss, while at the same time the system faces a fast-building wave of teacher retirements (Van der Berg et al. 2020). Given the renewed fervour with which expectations of ed-tech have risen following Covid-19-related school disruptions, as well as the need to address learning loss at scale in ways that do not require additional educators, the need for rigorous evaluations of ed-tech innovations is urgent.

3 Click Learning's implementation model

Click deploys the Reading Eggs suite of online English literacy applications in underprivileged primary schools across South Africa. Learners in Grades R to 7 work at their own pace through the applications, and with a programme-participation target of 22 hours per learner per year, Click hopes to improve learners' English literacy.

The first step for any ed-tech intervention is working, reliably-connected hardware. This is notoriously difficult to achieve in low-income contexts (Global Education Evidence Advisory Panel 2020). Most of the Click schools had no ICT infrastructure prior to intervention, and this section details the factors leading to effective technical implementation and support.

Click works closely with schools to identify suitable venues for laboratories, which can include staff rooms, classrooms, storage rooms, and libraries. In the absence of a suitable space, mobile tablet trolleys are deployed directly into classrooms. Of Click's 224 school laboratories, 92% are computer-labs and 7.5% are tablet-trollies. Schools provide various resources (e.g. classroom tables, chairs, security, and air conditioners) and cover power and general lab-readiness costs (approximately 22% of total laboratory setup costs), while Click installs the equipment.

Effective ed-tech interventions are highly reliant on functioning devices, quick-response troubleshooting, technical support and connectivity maintenance. In surveys with school principals in the Western Cape, Naicker (2013) found that the biggest perceived constraints to the use of computers in schools were procurement, maintenance, and support. Most ed-tech interventions in South Africa do not take on the responsibility for hardware, software, and connectivity procurement and maintenance in schools, and it is often at this first hurdle that they fail, as without maximising learners' time using functioning ICT, programme effectiveness is at high risk.

While there is some variation across sites, Click's general laboratory setup comprises an average of 45 devices (desktop computers, laptops and/or tablets). Since 2012, Click has bought 8,540 devices (of which approximately 6,830 are still in use), and manages, insures, and maintains an additional 5,270 devices owned by schools.

Click has a network of suppliers for installations, repairs, and replacements, allowing for quick resolution and turnaround, and employs centralised warehousing, courier services and private satellite repair teams. The use of regular audits, remote monitoring tools, ticketing systems and asset management software ensures that devices are maintained, fixed and replaced and downtime is limited. Click aims to get devices repaired within 48 hours of being reported and creative solutions are employed to further device life span (for example, faulty devices are scrapped for spare parts).

The programme relies on a constant data stream to deliver and monitor application usage; there is no offline version of Reading Eggs. Click partners with Rain as its primary internet service provider; the company provides pro bono data for 195 of the 224 schools.

Security is one of the greatest challenges, and 6% of devices were stolen in 2019. In 2020, substantial security investments were made into hidden remote monitoring cameras and safes for devices. A new set of rigorous Standard Operating Procedures (SOPs) was implemented, and Click collaborated with the South African Police Service, which led to arrests. Device theft subsequently dropped to 1% in 2021.

3.1 Systems and software

The Reading Eggs suite of online literacy applications delivers learner-responsive, level-appropriate English content designed to build the five component skills for successful reading as identified by the National Reading Panel (NICHD 2000): phonological awareness, decoding, vocabulary, oral reading fluency, and comprehension. The applications feature games, videos, e-books, and quizzes, as well as animated characters and songs, and learners are awarded tokens and prizes as they make their way through increasingly sophisticated levels of activity at their own pace. The platform's ability to track progress and match activities to each learner's ability is especially helpful in the South African classroom context, with learning losses (Ardington et al. 2021) exacerbating the already-challenging heterogeneity within classes.

Table 1 Suite of Reading Eggs early literacy programmes

E-learning Application	Grades	Focus Area
Fast Phonics	R-1	Covers phonics skills including letter-sound recognition, blending and spelling
Reading eggs	2-3 (4 for catch-up)	Covers phonics and phonemic awareness, sight words, vocabulary, comprehension and reading for meaning
Reading Eggspress	4-7	Builds reading, comprehension, spelling, vocabulary, and grammar

Click's implementation and scale is supported by a platform that manages and tracks end-to-end delivery, from early design and school selection to implementation and monitoring. The platform extracts learner-level data from the government's learner management systems, the South African School Administration and Management System (SA-SAMS) (via the Data-Driven Districts platform (DDD)), and from the Centralised Education Management Information System (CEMIS), used by the Western Cape Education Department (WCED). This information is used to create a unique login for each learner, and track their progress.

In order to participate in the Click programme, schools need to be no-fee or low-fee public primary schools, have strong buy-in from principals, and be located within internet connectivity coverage. Schools are responsible for reaching and maintaining stipulated usage targets, managing facilitators, updating the school's learner-management system and implementing the SOPs.

Click delivers the Reading Eggs suite to EFAL and English HL learners in Grades R to 7. Schools commit to ensuring that learners participate for a minimum of 22 hours per annum, and generally allocate a minimum of between one and two hours of lesson time per week.

Usage is monitored weekly, and Click tracks:

1. Numbers and percentages (active/registered) of learners using the programme at a class level, with a target 80% of learners; and
2. Length of time for which learners used the programme, with a minimum target of 40 minutes per week.

Click engages over 580 youth facilitators to deliver the intervention and ensure that laboratories are running optimally. Their responsibilities include opening and closing laboratories, assisting learners to login, and reporting any issues. Two facilitators are allocated to each computer laboratory or tablet-trolley, and are based at the school for an average of seven hours per day. The recruitment process involves identifying 18- to 26-year-old South Africans with a matric pass in English and maths or maths literacy, who are passionate about technology and working with children, and have basic computer skills. They undergo extensive training, with ongoing professional development. The facilitators receive a stipend equivalent to the national minimum wage.

High capital outlays on infrastructure are a major barrier to the cost-effectiveness of most ed-tech interventions. Chuang et al. (2021) identify the following recurrent, often hidden costs in implementing ed-tech: network equipment and server hardware (upgrades and spare parts), costs of replacing or repairing hardware, licence fees and software costs, cost of piloting, testing and customising software, energy usage, connectivity, increased security and school building infrastructure, support team salaries, and indirect labour costs.

Click's current annual infrastructure cost per learner is R65, which includes the costs of learner devices, peripherals (mice and headphones), connectivity infrastructure and laboratory security, amortised over three years. Click's annual budget for 2022, reaching more than 200,000 full-year-equivalent learners has implementation costs per learner at R305. This excludes the cost of facilitators (which adds a further R130 per learner). The total programme cost-per-learner per annum is thus R500.

Click partners with schools, government departments, foundations, and the private sector to implement its programmes. Provincial-level government department memorandums of understanding provide a mandate to operate in schools, and Click works with district officials to select schools, share resources (human and financial), support programme implementation and mediate any disputes. Schools contribute towards the costs of software licences, facilitators, security, and hardware. In 2021 schools contributed 7% of the total programme costs. Forty-two per cent of Click facilitators are funded by the YES programme (which covers their first year), while corporates, schools, and implementing partners cover the rest.

Click is primarily funded by South African family and corporate foundations, as well as companies' corporate social investment (CSI) and skills development funding, and an annual fundraising event, which together cover most of the annual budget.

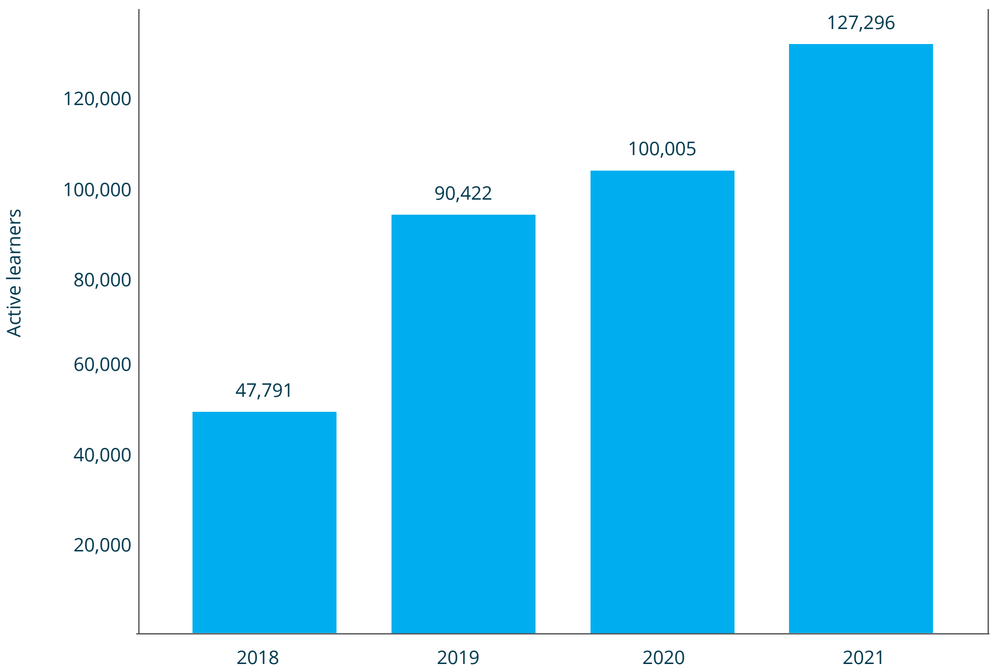
4 Monitoring, evaluation and learning

Click collects data on individual learner environments, usage and progression, and is currently further developing its monitoring, evaluation, and learning (MEL) systems and structures to measure impact on learning outcomes. This section describes the data, plans for MEL, and possibilities for further research.

4.1 Data description

A primary source of data for ongoing MEL is ClickMart, the large data warehouse that gathers information on learners and their schools, learner progress through the programme and their literacy-assessment results (measured through an eQuiz). Between 2018 and 2021 there were 184,710 learners in the sample, who were observed 365,514 times on the applications (many learners used the applications over more than one year). We restricted our analysis to active learners (those using the application at least once in that year), and Figure 2 below presents Click's growth in total active learners over time.

Figure 2 Total active learners, 2018-2021

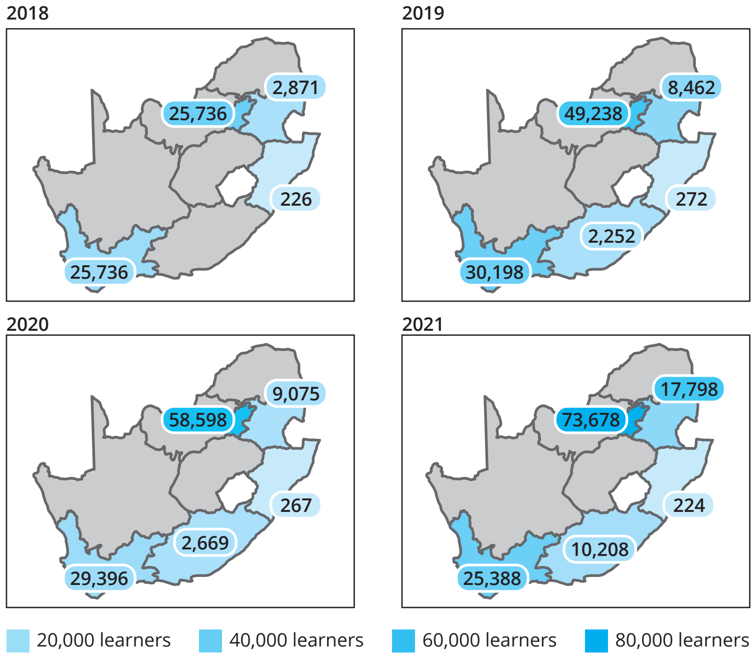


Note: *n = 365,514 observations of 184,710 individuals*

Around 90% of active Click learners are between Grades R and 4, with the remaining 10% between Grades 5 and 7. Between 2018 and 2021, the learner base aged slightly, with Grades 3-4 moderately more prominent compared to Grades R-2.

Click delivers programming in five provinces, illustrated in Figure 3. By 2021, the number of learners had grown to nearly 130,000. Of these, 58% were in Gauteng, 20% in the Western Cape, 14% in Mpumalanga, 8% in the Eastern Cape and the remainder in KwaZulu-Natal.

Figure 3 Click learners per province, by year



The Click learner group is by no means homogenous, with demographics such as mother-tongue, ethnicity, and English-subject changing as Click expands its geographical reach. Click learners are largely South African citizens (95%) and programme participation is evenly distributed between girls and boys. About 65% of learners between 2018 and 2021 took EFAL, whereas just over 28% took English HL[2].

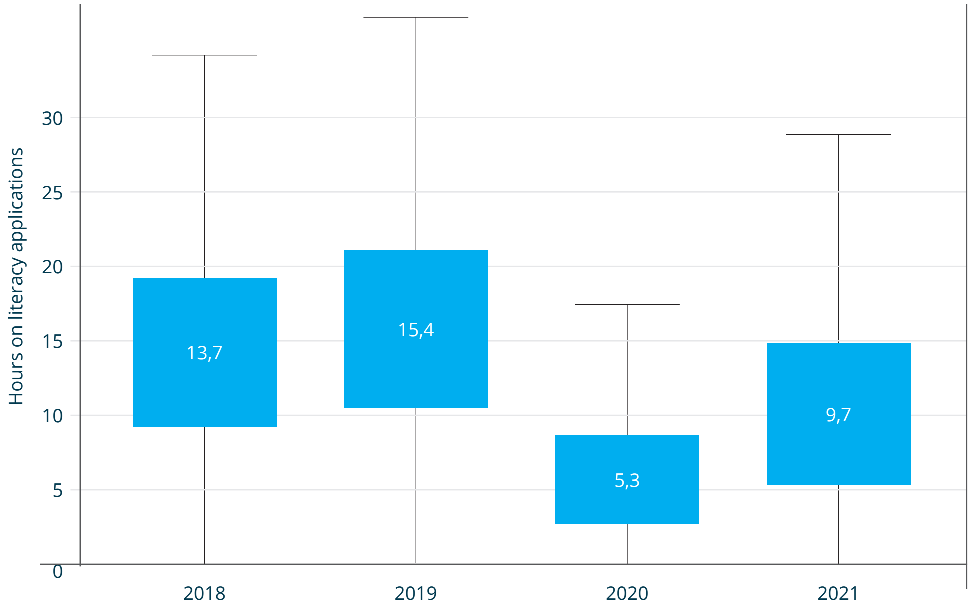
4.2 Usage

There is some way to go in reaching the targeted 22 hours per year. Figure 4 is a box and whisker plot in which learners are placed in order from least to most hours of use. Usage in 2018 and 2019 was clustered in ten-hour bands around the medians (13.7 hours and 15.4 hours respectively), where 50% of learners used the applications for an average of between ten and 20 hours (indicated by distribution of the coloured boxes). The whiskers on either side of the boxes represent the approximately 25% of learners with the lowest usage and the 25% of learners with the highest usage. The number of higher-frequency users increased

2 The remaining 6.3% of learners either took English HL or EFAL but it was not possible to discern which due to their enrolment in parallel-medium schools offering both subjects.

over time: in 2018 roughly 25% of learners used the applications for between 20 and 35 hours, relative to 2019 where 25% used the applications for between 20 and 40 hours. The progress made in 2019 towards achieving 22 usage hours was thwarted by Covid-19.

Figure 4 *Reading Eggs* **suite of applications usage by year**



Note: *n = 365,514. Medians labelled with values.*

The disruptions of 2020 saw the median usage fall to 5.3 hours per year, with far less variability across learners. In 2020, 50% of learners were using the applications for between 2.5 and 8 hours per year and the top 25% of users were only engaging for between 8 and 17.5 hours per year. In 2021 usage improved, with 50% of learners engaging for between 5 and 15 hours, and the top 25% for between 15 and 27.5 hours.

Table 2 presents the percentage of learners spending a threshold level of hours on the applications per year for 2018 and 2019 (2020 and 2021 are excluded from the table given the disrupted school year.) It is encouraging to see the improvement at every threshold from 2018 and 2019 withstanding the significant increase in learner numbers.

Table 2 **Percentage of learners spending greater than X hours annually on the Reading Eggs suite of literacy apps, 2018 and 2019**

Year	5 hours	10 hours	15 hours	20 hours	N
2018	92.3%	71.5%	42.8%	22.8%	47,791
2019	93.4%	77.7%	52.2%	28.7%	90,422

Research is needed to determine the optimal dosage for achieving learning gains. The stated goal of each learner spending at least 22 hours on the applications per year is driven by the available English subject slots in the school timetable. Click plans to commission research to determine the number of hours required to achieve minimum scores in foundations of literacy and comprehension respectively.

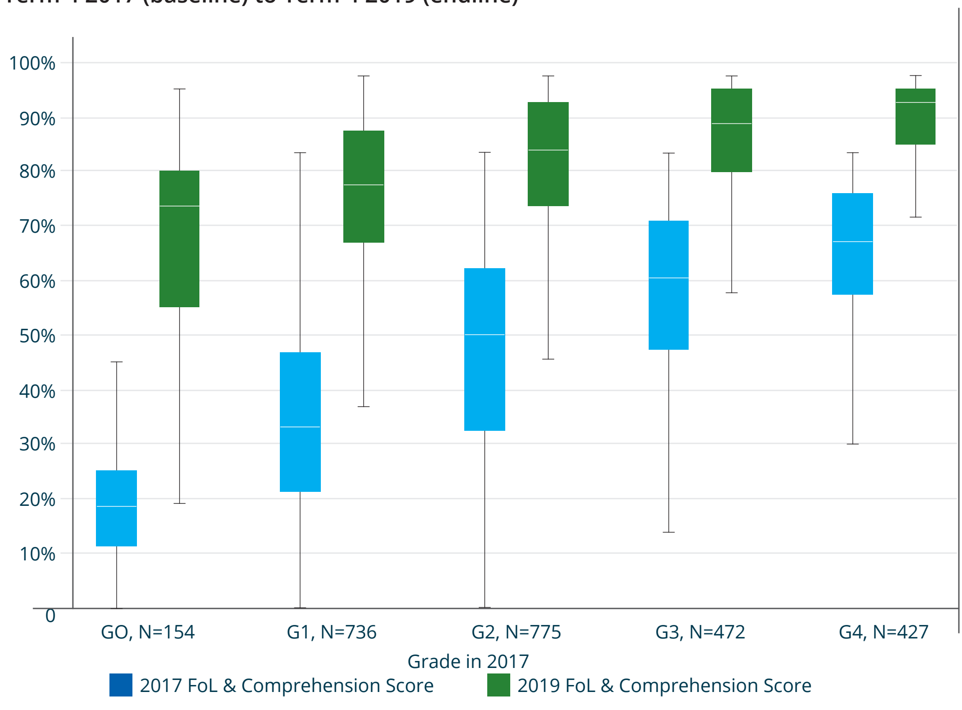
4.3 Impact analysis roadmap

Click has developed an eQuiz which is based on the EGRA, and assesses phonological awareness (knowledge of letter-sound relationships), decoding (sounding-out words), word recognition, vocabulary, oral reading fluency (speed and accuracy), and comprehension. The eQuiz can be conducted quickly and at minimal cost, and results are availed in real-time. This presents opportunity for literacy assessment and research in South Africa, given the large expense and amount of time required to conduct paper-based assessments. The eQuiz allows for baseline measurement (once learners have had a week or two to learn basic computer use), and can be conducted at various points throughout the intervention.

Click contracted a data analytics firm to review the eQuiz data and make recommendations. Given the disruptions of Covid-19 from 2020, data from between the ends of 2017 and 2019 was analysed. Learners completed an assessment at the end of 2017, and again at the end of 2019. Literacy levels were measured on a subset of approximately 2,000 learners that answered common questions of the eQuiz in both instances.

Preliminary results are presented in the box and whisker plot in Figure 5 with the light pink boxes indicating results from 2017 and the dark pink from 2019. A substantial improvement in literacy over this time period was seen for all grades. For instance, Grade Rs in 2017 had a median of 0.19 and 50% scored between 0.12 and 0.25. By 2019, Grade Rs had a median of 0.75 and 50% scored between 0.58 and 0.8.

Figure 5 **Box and whisker plots of literacy gains by grade for subset of Click learners, from Term 4 2017 (baseline) to Term 4 2019 (endline)**

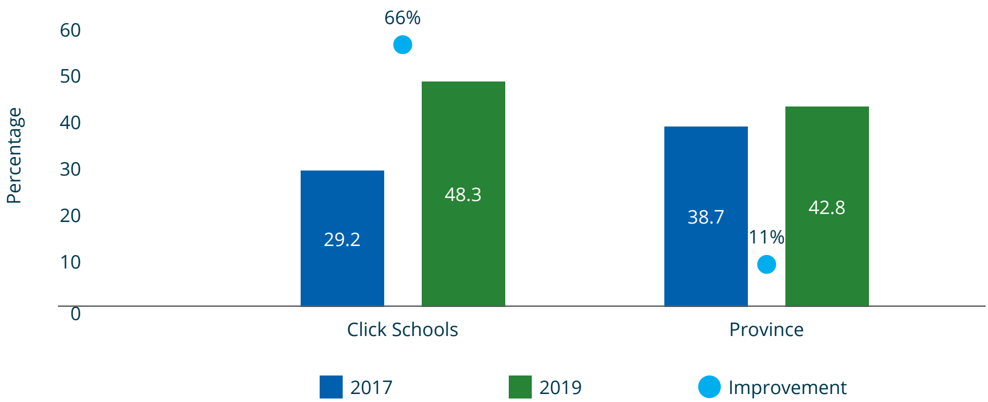


Source: Click internal data analysed by independent M&E consultants

Given that literacy gains are expected over a period of two years independent of any intervention, it is useful to compare the Click learners to a benchmark group that were not using the applications. The WCED conducts systemic assessments with learners in Grades 3, 6 and 9. We conducted a preliminary analysis of the Grade 6 English systemic test conducted by the WCED in 2019. Using the 2017 Grade 6 cohort as a quasi-control, we then compared the difference in learners' performance (those in Click schools versus the provincial average) with that of the 2019 cohort.

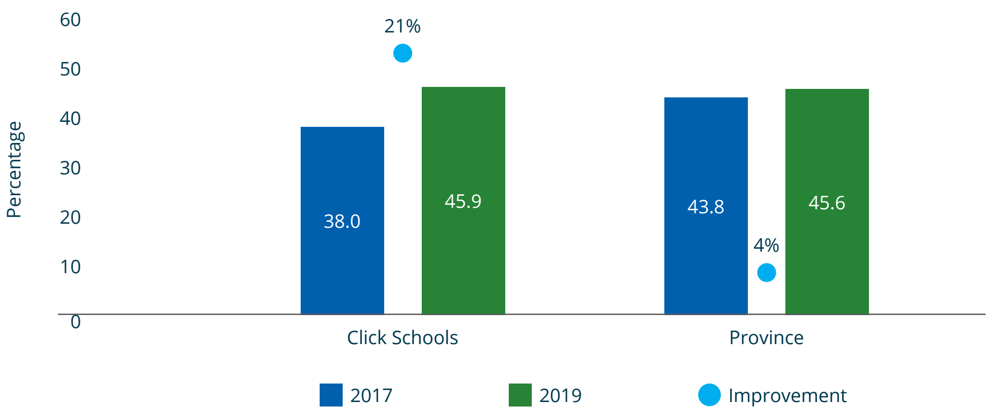
The analysis finds that from a lower base, learners participating in the Click programme for an average of 17.5 hours per year significantly outperformed the provincial averages. While this does not present evidence of impact due to the very small sample size (2,563 learners in four schools) and potential sampling biases, it is an encouraging indication that warrants further exploration.

Figure 6 Percentage of Grade 6 learners passing the Systemic Tests in 2017 and 2019 in four Click schools vs provincial results.



Source: Provincial systemics data from WCED (2019), analysis by Click Learning

Figure 7 Average Grade 6 marks in the Systemic Tests in 2017 and 2019 in four Click schools vs provincial results.



Source: Provincial systemics data from WCED (2019), analysis by Click Learning

4.4 Next steps

The eQuiz will be run alongside the paper-based EGRA with a subsample of learners, to validate the tool. Once this has been achieved, Click has longitudinal per-learner data on programme usage and literacy levels, as well as demographic and environment information which can be analysed to better understand the programme's impact. Dosage is the first area of focus, with a view to determining the threshold level of hours required to effect gains for learners with different literacy levels, and in different grades.

Once a sufficiently large sample size is available, Click will be commissioning research to employ the eQuiz / EGRA data and a within-school, fixed-effects framework. The study will look at a particular grade within a school over time, and compare the performance of different cohorts based on the assumption that learners in a particular grade are a good counterfactual for learners in the same school in that grade two years later (with the requirement that the years do not straddle the onset of school closures and rotational timetables).

Going forward, Click's planned scale for 2024 creates interesting possibilities for evaluation, as it will be working with 600 schools, some of which could be randomly selected from a list of eligible schools.

5 Conclusion and way forward

Chuang et al. (2021) note Adam et al.'s (2018) non-exhaustive list of questions to inform readiness for large-scale ed-tech adoption:

- "Is the digital infrastructure mature enough?
- What available devices can be leveraged for learning?
- Is there sufficient human capacity, in addition to opportunities for continuous professional development?
- Is there a supportive ecosystem across the developer community, donor community, and education community?
- Is there sufficient funding to develop, adapt, and maintain the initiative?
- Is there support from political economy stakeholders, policy and data systems?"

Click has built a programme that, for the most part, can answer 'yes' to each of these, and in doing so, has learned valuable lessons.

In order to recover from missed schooling and learning losses due to the pandemic, there is an urgent need for additional support to learners for catch-up. The potential that interventions such as Click presents for large-scale learner support needs to be investigated. The big research questions in this regard are (1) What is the programmes' causal impact on literacy and numeracy outcomes? and (2) What is the required minimum dosage? Research is needed to establish whether level-appropriate, self-led ed-tech programmes can improve learners' literacy levels, fill gaps and regain learning losses. And if so, how these interventions can be scaled as cost-effectively and efficiently as possible.

References

Ardington, C., Wills, G., & Kotze, J. (2021). Covid-19 learning losses: Early grade reading in South Africa. *International Journal of Educational Development 86*. 102480.

Banerjee, A., Cole. S., Duflo, E., & Linden, L. (2005). Remedying Education: Evidence from two randomized experiments in India. NBER Working Paper No. 11904. National Bureau of Economic Research, Cambridge, MA.

Berry, J. & Mukherjee, P. (2016). Pricing of Private Education in Urban India: Demand, Use and Impact. Unpublished.

Böhmer, B., Burns, J., & Crowley, L. (2014). Testing Numeric: Evidence from a randomized controlled trial of a computer based mathematics intervention in Cape Town high schools, (Master's thesis) University of Cape Town, Cape Town.

Castillo, N. M. (2017). Technology for improving early reading in multilingual settings: Evidence from rural South Africa. Publicly Accessible Penn Dissertations. 2204.

Cheung, A. C. K., & Slavin, R. E. (2012). How features of educational technology applications affect student reading outcomes: A meta-analysis. *Educational Research Review* 7, 198–215.

Cheung, A. C. K., & Slavin, R. E. (2013). Effects of educational technology on reading outcomes for struggling readers: A best-evidence synthesis. *Reading Research Quarterly, 48*(3), 277–299.

Chigona, A., Chigona, W., Kayongo, P., & Kausa, M. (2010). An empirical survey on domestication of ICT in schools in disadvantaged communities in South Africa. *International Journal of Education and Development using Information and Communications Technology (IJEDICT)* 6(2), 21–32.

Chuang, R., Burnett, N., & Robinson, E. (2021). Cost-Effectiveness and EdTech: Considerations and case studies. (Helpdesk Response 32). EdTech Hub.

Combrinck, C., & Mtsatse, N. (2019). Reading on paper or reading digitally? Reflections and implications of ePIRLS 2016 in South Africa. *South African Journal of Education*. *39, Supplement* 2, S1–S14.

Comings, J. (2018). *Assessing the impact of literacy learning games for Syrian refugee children: An executive overview of Antura and the letters and Feed the Monster impact evaluations.* World Vision, Inc. and Foundation for Information Technology Education and Development, Inc.

Comings, J. P. (2020). A case study of innovations that address constraints on the effectiveness of early grade reading achievement. *PROSPECTS*. Springer.

Department of Basic Education. (2020). Portfolio Committee Presentation: ICT roll-out. Available at: https://pmg.org.za/committee-meeting/30052/ (Accessed 22 July 2022).

Department of Basic Education. (2019). National Education Infrastructure Management System (NEIMS) Report.

Global Education Evidence Advisory Panel. (2020). Cost-effective approaches to improve global learning: What does recent evidence tell us are 'Smart Buys' for improving learning in low- and middle-income countries? FCDO, UK & World Bank.

Grand View Research. (2022). Education Technology Market Size, Share & Trends Analysis Report, 2022–2030.

Howie, S. J., Combrinck, C., Roux, K., Tshele, M., Mokoena, G.M., & McLeod Palane, N. (2017). PIRLS LITERACY 2016: South African Highlights Report. Pretoria: Centre for Evaluation and Assessment.

Howie, S. J., Combrinck, C., Roux, K., Tshele, M., Mtsatse, N., McLeod Palane, N., & Mokoena, G.M. (2017b). ePIRLS 2016: South African Highlights Report. Pretoria: Centre for Evaluation and Assessment.

Karsenti, T., Collin, S. & Harper-Merrett, T. (2012). *Pedagogical Integration of ICT: Successes and Challenges from 100+ African Schools* Ottawa, ON: IDRC.

Lesnick, J., George, R. M., Smithgall, C., & Gwynne, J. (2010). *Reading on grade level in third grade: How is it related to high school performance and college enrolment*? Chicago: Chapin Hall at the University of Chicago.

Ma, Y., Fairlie, R., Loyalka, P., & Rozelle, S. (2020). *Isolating the 'tech' from edtech: experimental evidence on computer-assisted learning*. National Bureau of Economic Research (NBER), Working paper 26953.

Muralidharan, K., Singh, A., & Ganimian, A. (2018). Disrupting Education? Experimental evidence on technology-aided instruction in India. *American Economic Review*, 1–94.

Naicker, V. (2013). Understanding the role of principals and the use of computers in secondary schools. *South African Journal of Higher Education 27(4)*, 966–984.

National Institute of Child Health and Human Development (NICHD). (2000). Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups. *NIH Publication No. 00-4754*. Washington, DC: NICHD.

Ndlovu, N., Maringe, F., & Chiramba, O. (2021). Learning with technology in low-income households in times of disruption. Zenex Foundation, Pretoria.

Perera, M., & Diego, A. (2017). Diferencias por Género y Contexto Socioeconómico del Impacto de la Plataforma Adaptativa de Matemática. CINVE.

Piper, B., Zuilkowski, 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.

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.

Reddy, V., Winnaar, L., Juan, A., Arends, F., Harvey, J., Hannan, S., Namome, C., & Zulu, N. (2020). *TIMSS 2019 Highlights of South African Grade 5 Results in Mathematics and Science*. HSRC. Pretoria. Available at: http://www.hsrc.ac.za/uploads/pageContent/1044991/TIMSS%20 2019_Grade%205_HSRC_FinalReport.pdf (Accessed 21 February 2022).

Reddy, V., Winnaar, L., Juan, A., Arends, F., Harvey, J., Hannan, S., Namome, C., Sekhejane, P., & Zulu, N. (2020a). *TIMSS 2019 Highlights of South African Grade 9 Results in Mathematics and Science*. HSRC. Pretoria.

Rodriguez-Segura, D. (2021). Educational technology in developing countries: a review of the evidence. *The World Bank Research Observer* 37(2), 171–203.

Skryabin, M., Zhang, J., Liu, L., & Zhang, D. (2015). How the ICT development level and usage influence student achievement in reading, mathematics and science. *Computers and Education 85*, 49–58.

Smith, G. S. & Hardman, J. (2014). The Impact of Computer and Mathematics Software Usage on Performance of School Leavers in the Western Cape Province of South Africa: A Comparative Analysis. *International Journal of Education and Development using Information and Communication Technology 10*(1), 22–40.

United Nations. (2019). *Value Creation and Capture: Implications for developing countries*. Digital Economy Report. United Nations Conference on Trade and Development.

Van der Berg, S., Gustafsson, M., & Burger, C. (2020). School teacher supply and demand in South Africa in 2019 and beyond. Department of Higher Education and Training, Pretoria.

Vegas, E., Ziegler, L., & Zerbino, N. (2019). *How ed-tech can help leapfrog progress in education*. Centre for Universal Education, Brookings Institute.

Vromant, N., Kuppens, L., Hazemba, M., Mwamba, F., & Cupito, E. (2021). Scaling Teaching at the Right Level: Insights from the Catch Up Literacy and Numeracy Intervention in Zambia. In: Munene, I. (Ed). *Ensuring All Children Learn: Lessons from the South on What Works in Equity and Inclusion*. Lexington Books, United Kingdom.

10

The Grade R Mathematics Project (R-Maths) in the Western Cape 2016-2019: Conceptualisation, development, implementation, and evaluation

GARTH SPENCER-SMITH, ELEANOR HAZELL & CALLY KUHNE

Abstract

The Grade R Mathematics Project (2016-2019) was a province-wide intervention in the Western Cape, led by the Western Cape Education Department, and aimed at improving the quality of Grade R mathematics teaching and learning in the province. Content materials and support were provided to Foundation Phase subject advisors by the Schools Development Unit at the University of Cape Town. The SAs trained all Grade R teachers in the province in phases. The training of SAs and the first phase of Grade R teacher training was evaluated, as was the consolidation phase.

Change in Grade R mathematics knowledge and understanding of SAs and Grade R teachers was measured via a non-experimental pre-test-post-test evaluation design. Change in teaching practice was assessed via a baseline and endline observation of Grade R mathematics lessons in six case-study schools. Change in the Grade R learners' mathematical knowledge and skills was assessed via a quasi-experimental difference-in-difference design, with the learners selected randomly from two districts. During Phase 1 of the intervention (2017), 1,439 teachers were trained, allowing the 1,997 teachers allocated to the second implementation phase to be regarded as a comparison group. The

KEYWORDS
mathematics, Grade R, provincial programme, systemic change

learning gains of a random selection of learners in the two groups (Phase 1 intervention and comparison) were then compared.

The evaluation showed that the impact on both the SA and teacher knowledge was significant, with a large effect size. The case study findings indicated that teachers used the R-Maths Programme materials to plan and implement lessons, whilst the recommended lesson structure was broadly adhered to. However, classroom management did not appear to change substantially. The impact of R-Maths on the Grade R learners' mathematical knowledge and understanding after eight months was small but positive. The results of a general linear model showed that a significant factor on learner performance was the group the learner was in (i.e. intervention or comparison), with the intervention group performing better than the comparison group, with a small effect size.

The chapter concludes with recommendations based on lessons learned from the implementation and evaluation of the R-Maths Project that should ideally be enacted in any large-scale school intervention.

1 Introduction

Within the context of early grade learning, play-based pedagogies have received considerable attention recently, both in South Africa (Centre for the Developing Child 2016; UNICEF, 2020) and elsewhere (Whitebread et al. 2017; Sylva et al. 2007; Zosh et al. 2017). There is broad agreement that early-childhood programmes require a well-rounded curriculum, active play with concrete materials, mediated adult/child interaction with groups as well as child-initiated individual activities. While there are various national policy commitments to adopt such approaches in the early years (DBE 2011 & 2015), there are very few examples of how this can be realised in practice in early grade classrooms. This chapter presents insights from a province-wide intervention to improve mathematics teaching and learning in Grade R in the Western Cape.

In 2016, the University of Cape Town's (UCT) Schools Development Unit (SDU), in partnership with the Western Cape Education Department (WCED), developed the Grade R Mathematics Programme (R-Maths Programme). This programme consists of training materials developed to strengthen the learning and teaching of mathematics in South African Grade R settings and to support teachers to implement the Curriculum Assessment Policy Statement (CAPS)[1] in WCED public schools and registered pre-schools or crèches. The programme is specifically designed to provide teachers with the mathematics knowledge and understanding of the 'big ideas' and concepts that underpin the CAPS content and appropriate classroom activities and methodology to implement this content.

The WCED concurrently initiated R-Maths to train 77 Foundation Phase (FP)[2]

1 These are the policy documents produced for each subject and grade by the Department of Basic Education (DBE). They include, inter alia, a detailed syllabus and assessment policy.

2 Formal schooling in South Africa begins with the Foundation Phase and includes four years (Grade R to Grade 3). Children enter Grade R at 5 years turning 6. The Curriculum Assessment and Policy Statements (CAPS) provide content for the subjects offered at this phase: Home Language, First and Second Additional Language (Grades 1–3 only), Life Skills and Mathematics.

subject advisors (SAs) and 3,200 Grade R teachers[3] in all education districts in the Western Cape province in the use of the R-Maths training and classroom materials.

R-Maths had five stages: (a) materials development; (b) SA training; (c) Phase 1 of Grade R teacher training; (d) Phase 2 of Grade R teacher training); and (e) the consolidation phase. The development of the mathematics programme materials for the use of trainers and in the classroom was concluded in 2016. SAs were first trained over five days in late 2016. The following year, in Phase 1, roughly half of Grade R teachers in the Western Cape were trained over a six-month period by means of seven two-hour workshops at sub-district level (cluster trainings (CTs)) and one five-day block training. In 2018, as part of Phase 2, the remainder of the Western Cape Grade R teachers were trained over a six-month period. This accounted for a total of about 3,200 trained teachers. The WCED determined that it was not viable to train all Grade R teachers at once, so each education district divided Grade R teachers into two cohorts. Districts used different criteria to determine which teachers would be trained in cohorts 1 and 2: some divided teachers into cohorts based on geography and others on the basis of school performance (e.g. starting with better-performing or worse-performing schools).

In 2019, during the consolidation phase, approximately 600 novice teachers were trained through 12 one-day workshops.

Figure 1 provides a brief overview of the timelines of the R-Maths implementation:

Figure 1 Timelines of R-Maths implementation

2016	2017	2018	2019
Materials development **Subject advisor training**	**Phase 1** 7 × 2 hr cluster workshops 5-day block training	**Phase 2** 7 × 2 hr cluster workshops 4-day block training	**Consolidation** 12 × 6 hr training, novice teachers by lead teachers

R-Maths was evaluated from 2016 to 2018 by a partnership of two evaluation agencies: JET Education Services (JET) and Kelello Consulting. Kelello also evaluated the consolidation phase in early 2020. The first evaluation was comprehensive, involving clarification of R-Maths theory[4], assessment of implementation, a pre- and post-assessment of SAs' and Grade R teachers' knowledge of Grade R mathematics teaching, case studies of Grade R teaching in schools and a quasi-experimental impact assessment comparing the mathematical knowledge and skills of Grade R learners whose teachers had and had not participated in the R-Maths Project. The evaluation focused on two districts in the Western Cape, one rural and one urban.

Fleisch (2018) refers to an "educational triple cocktail" that consists of lesson plans, quality learning materials, and one-on-one coaching. This cocktail has been shown to be effective in improving teaching, and ultimately, learner performance

3 The persons teaching Grade Rs are referred to throughout this chapter as 'teachers'. It is acknowledged that most of the individuals teaching Grade R in the Western Cape and elsewhere in South Africa tend to hold an NQF level 5 certificate, rather than a bachelor's degree, and thus are referred to by education departments as 'practitioners'. We refer to these individuals as 'teachers' rather than the more unwieldy 'teachers/practitioners'.

4 This presents "a plausible and sensible model of how a programme is supposed to work" (Bickman 1987, 5). It is different from Science and Social Science theories, which explain facts or phenomena. Programme theory includes a theory of implementation (how the project will be implemented) and a theory of change (how the project will bring about change).

in South Africa (Fleisch, 2018; Fleisch & Motilal, 2020; Kotze, Fleisch & Taylor, 2019). R-Maths enacted this cocktail with some deviations. In particular, R-Maths made use of a modified-cascade model to train and support teachers. WCED FP SAs were trained and supported to train Grade R teachers. Known weaknesses of this model are that key messages get distorted and diluted at each level of the cascade (McDevitt 1998; Suzuki 2008; Weddell 2005). However, the SAs were themselves supported to train and support teachers and a benefit of this approach was that the intervention built the capacity of existing human resources, thus also supporting embedding and sustainability. There is some evidence from elsewhere in Africa (see Gove et al. 2017) that a model, which works with and through existing education system human resources can be effective at improving learner performance.

2 Literature

A play-based instructional programme is informed by the idea of a continuum of different types of play, with optimal learning taking place when the activity is joyful, meaningful, actively engaging, and involves iterative thinking and social interaction (Jensen et al. 2019; Montie et al. 2006; Pyle et al. 2017; UNICEF, 2020; Whitebread et al. 2017; Whitebread et al. 2015; Zosh et al. 2018).

The motivation for the design of this R-Maths instructional programme was premised on the belief that teachers do not share a common understanding of children's mathematical development and the complexity and diversity of this process (Kuhne et al. 2005). The R-Maths Programme was informed by current research on early mathematics and approaches to improving teachers' knowledge of this development using learning 'pathways' or trajectories (Clements & Sarama 2014; Sarama & Clements 2009; Sztajn & Wilson 2019; Van den Heuvel-Panhuizen et al. 2012). These pathways provide steps in learning a specific topic and each new step in the learning path builds on the earlier steps (National Research Council, 2009). Learning pathways provide structure and explicit content that is required to plan appropriate instructional tasks, with specific goals, so that the learner progresses through a developmental continuum of thinking in a particular mathematical domain (Kuhne et al. 2013).

3 Design and data

3.1 Research questions

The following research questions will be answered in this chapter:

1. What was the impact of R-Maths on FP SAs, Grade R teachers and Grade R learners?
2. What recommendations can be made for the roll-out of other early grade interventions from lessons learned implementing and evaluating R-Maths in the Western Cape?

3.2 Evaluation design[5]

Randomisation was not possible; thus, a quasi-experimental evaluation design was employed whereby a treatment group of Grade R learners whose teachers participated in Phase 1 of the project were compared to a group of Grade R learners whose teachers would participate only in Phase 2. Both groups were tested at baseline and endline and the difference-in-differences of their baseline and endline scores was compared. This use of this method controls for differences between the groups at baseline, resulting in a more accurate measurement of impact (White & Sabarwal, 2014).

Change in the Grade R-level mathematics knowledge was measured via a pre-test before training, and a post-test after training. All the FP SAs in the Western Cape and the Phase 1 teachers in the two focus districts (one urban, one rural) who gave their informed consent took these tests. Change in the Grade R teachers' teaching practice was assessed via a baseline and endline observation of eight teachers teaching Grade R mathematics lessons in six case-study schools (three schools in each of the focus districts).

A simple random sample (SRS) of learners from schools in the two focus districts was drawn. Learners whose teachers received training in Phase 1 of implementation were the intervention group, whilst learners whose teachers were only trained in Phase 2 formed the comparison group. All schools were included in the sampling frame for the pre-testing and post-testing. Then, the enrolment numbers in each school were used to allocate an identifier to each learner. For example, if school 'S' had 82 Grade R enrolments according to administrative data, learner identifiers from S1 to S82 were generated. SRS was then used to select learners as per the details in the table below, resulting in the number of sampled learners differing from school to school.

Table 1 Number of schools and learners sampled, by group

	Intervention			Comparison		
	No. of schools	No. of learners	Ave. learners per school	No. of schools	No. of learners	Ave. learners per school
Urban district	33	168	5.1	108	168	1.6
Rural district	40	168	4.2	28	168	6.0

Thereafter, sampled schools were contacted and class lists were matched against the SRS selections. The sampled learners were then identified on the school lists and assessed. The same learners were assessed at baseline and endline. The table below provides information on the number of learners tested in each group.

5 The evaluation of the R-Maths intervention has previously been reported on in Hazell, Spencer-Smith & Roberts (2019)

Table 2 Number of learners tested, by time and group

	No. of learners tested (baseline)	No. of learners tested (endline)	Attrition percentage
Urban district (treatment)	167	160	4.2%
Urban district (comparison)	161	157	2.5%
Rural district (treatment)	166	150	9.6%
Rural district (comparison)	168	155	7.7%

There was some attrition due to factors such as absence from school on the day of endline testing or learners moving schools. However, this is a relatively small proportion and less than the 10% of 'extra' learners that were deliberately originally sampled (and tested) at baseline. Thus, there was no 'top up' of the sample at endline with any additional learners.

The evaluation focus districts were identified after investigating how well-balanced the phase 1 and phase 2 learner cohorts were, based on: number of schools, number of Grade R teachers, number of learners, school quintile[6], language of learning and teaching (LoLT) and Grade 3 systemic test[7] mathematics results. The rural and urban districts with the most balanced cohorts were selected. A comparison of important metrics for the two focus districts selected is presented below.

Table 3 Comparison of cohorts per focus district

Metric	Urban district		Rural district	
	Cohort 1	Cohort 2	Cohort 1	Cohort 2
Mean score Grade 3 systemic test mathematics 2014	69.5	61.1	49.7	45.4
Mean score Grade 3 systemic test mathematics 2015	73.2	65.8	46.8	44.1
Gain score 2014-2015	3.7	4.7	−2.8	−1.3
Q1-3 schools	6 (18%)	26 (24%)	32 (80%)	23 (82%)
Q4-5 schools	27 (82%)	82 (76%)	8 (20%)	5 (18%)
Afrikaans LoLT	2	3	27	23
English LoLT	14	38		
isiXhosa LoLT	6	23		
Afrikaans/Eng LoLT	11	44	5	5
Afrikaans/Eng/isiXhosa LoLT			8	0

6 In South Africa, all public schools are placed in a quintile. Schools in the highest quintile (5) are located in communities that have the highest income; those in the lowest (1) in areas where the communities have the lowest income.

7 The WCED runs annual 'systemic tests' in home-language and mathematics in selected grades (3, 6 and 9). These tests are written by all learners at public schools in the province.

3.3 Data collection

The table below indicates the data collection[8] activities undertaken during the evaluation.

Table 4 Evaluation data collection activities

Activity	Description of activity
Stakeholder interviews	Conducted with key project stakeholders at base- and endline.
FP subject advisor test	30 minutes long, conducted at baseline and endline; the test included items assessing the mathematical knowledge necessary for supporting Grade R mathematics teaching.
Teacher test	Conducted at baseline and endline with Grade R teachers in two case- study districts. Similar in content to the FP subject advisor test.
Learner test	A 47-item demonstration version of the Marko-D test of mathematical competencies conducted with a sample of Grade R learners in two districts at baseline and endline. Baseline testing was in February/March 2017 and endline testing in October/ November 2017.
Observation of training, cluster trainings and dry-runs[9]	1. Observation of some of the training conducted by the SDU: SA block training 2. CT and block training dry-runs for the training of teachers. 3. Observation of some of the training conducted by SAs, with SDU support: CTs and block training of teachers.
Case studies	Conducted in three schools in each of the two focus districts at baseline (February/March 2017) and endline (October/November 2017): lesson observations and interviews were conducted with Grade R teachers and interviews were conducted with FP Department Heads (DHs) and SAs.
Monitoring fidelity (tracking dosage)	Analysis of attendance registers supplied by WCED, and support reports provided by SDU, for activities with SAs and teachers in two districts.

The R-Maths consolidation evaluation conducted in early 2020 was a brief qualitative evaluation of the 2019 consolidation phase. The aim was to assess the mechanisms used by the WCED to embed R-Maths into the provincial, district, and school-level functioning and to determine how embedded the intervention was.

The consolidation evaluation involved the following data collection activities:

1. Key stakeholder interviews (with 22 individuals from the WCED, schools, the SDU, and funders).
2. School visits at four case-study schools, two from a rural and two from an urban district. During these visits, interviews were conducted with the Principal, FP DoH and Grade R teacher, a mathematics lesson was observed, and teacher-provided learner work was reviewed.

8 Ethical clearance to undertake research was sought from and provided by the WCED and the University of Johannesburg.

9 'Dry runs' refer to support provided to the SAs by SDU before each Gr R teacher training. Their purpose was to help to ensure that the SAs were ready to conduct the training.

3.4 Data analysis

The stakeholder interviews, observations and case-study data were analysed qualitatively. For the interview data, a coding framework was developed based on pre-identified themes and new codes were added during analysis to accommodate emergent themes. Trained observers used a semi-structured instrument to take notes and photographs during observed sessions. Case-study data were captured via a semi-structured observation instrument and structured interview guide, with ratings reviewed and moderated[10] by senior researchers. Case-study findings per instrument were reviewed side-by-side. Baseline and endline case-study reports were compiled and the reports per school were compared using the constant comparison method.

The test data for SAs, teachers, and learners were analysed quantitatively. For the SA and teacher test data, descriptive statistics per group (such as means and standard deviations) were calculated for test-response data and represented as graphs and tables. Paired sample t-tests were used to compare the pre-test and post-test data and effect sizes were used to measure the sizes of the differences between the pre-test and post-test data. For the teacher-test data, this was done overall and per district.

The Marko D test was used to test learners. This test has an underlying theoretical framework pertaining to the cognitive development levels of children (Fritz et al. 2014; Fritz-Stratmann et al. 2014). The test assesses number-concept development, is appropriate for Grade R, and was administered orally and one-on-one in the LoLT of the school. Questions in the Marko D test are set at five different levels of conceptual understanding[11]. The English and Afrikaans versions used had been validated, but not the isiXhosa version used. The Marko D test was used in a slightly different way to the test designers' intent. Learners were always assessed in the LoLT of the school they attended[12]. This approach was applied consistently throughout the testing in both intervention and comparison group learners in rural and urban districts and thus provides a valid measure of the mathematics performance of learners in the LoLT of the school.

For the learner test, descriptive statistics were calculated per group, as indicated above. Independent sample t-tests were conducted to determine if the observed differences in the two districts were significant. A general linear model (GLM) was

10 Where the reviewer disagreed, the observer was asked to provide additional evidence to justify the rating and, in a few instances, ratings were changed by consensus.

11 Level 1: Counting (e.g. a knowledge of the sequence of number words; an ability to count with 1-1 correspondence);
Level 2: Mental number-line (e.g. a knowledge of preceding and succeeding numbers, given a number; ability to perform simple addition and subtraction calculations by 'counting all');
Level 3: Cardinality and decomposability (e.g. an understanding of the number of elements in a set or grouping);
Level 4: Class inclusion and embeddedness (e.g. an understanding that "each number can be decomposed into partial quantities which, together, are equivalent to the total quantity" (Fritz et al. 2014, 119); and
Level 5: Relationality (e.g. understanding that the intervals between any two successive numbers remains constant).

12 At dual- or parallel-medium schools, learners were assessed in the language in which they were being taught. For example, at a parallel medium Afrikaans-isiXhosa school, learners being taught in Afrikaans were assessed in Afrikaans, and vice versa.

developed which generated t-statistics and accounted for other factors or covariates (such as age, language, and quintile). This GLM was created for each level of the Marko D test and the Marko D test overall.

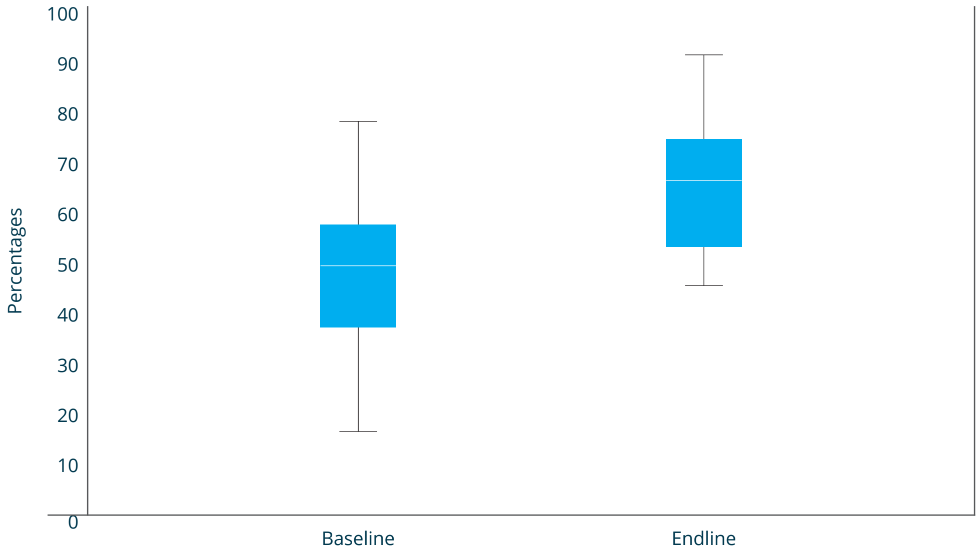
4 Findings

4.1 Impact on subject advisors

The evaluation showed that the impact on the SAs' Grade R pedagogical knowledge (PK), subject matter knowledge (SMK) and pedagogical content knowledge (PCK) was significant at the 95% confidence level (paired samples t-test, $p < 0.001$), with a large effect size[13] ($d = 1.44$). In the pre-test, the group mean was 48.6% and individuals obtained scores ranging from 16.7% to 79.2%. In the post-test, the group mean increased to 66.1% and individuals obtained scores that ranged from 45.8% to 91.7%. See Figure 2.

Test scores of SAs remained varied, indicating that their knowledge of supporting Grade R mathematics teaching was varied, and the quality of training and support provided to Grade R teachers was likely to be uneven.

Figure 2 FP SA's performance in the pre-test and post-test



13 All effect sizes provided in this chapter are Cohen's d-effect sizes. It is determined by calculating the mean difference between the two groups, and then dividing the result by the pooled standard deviation.

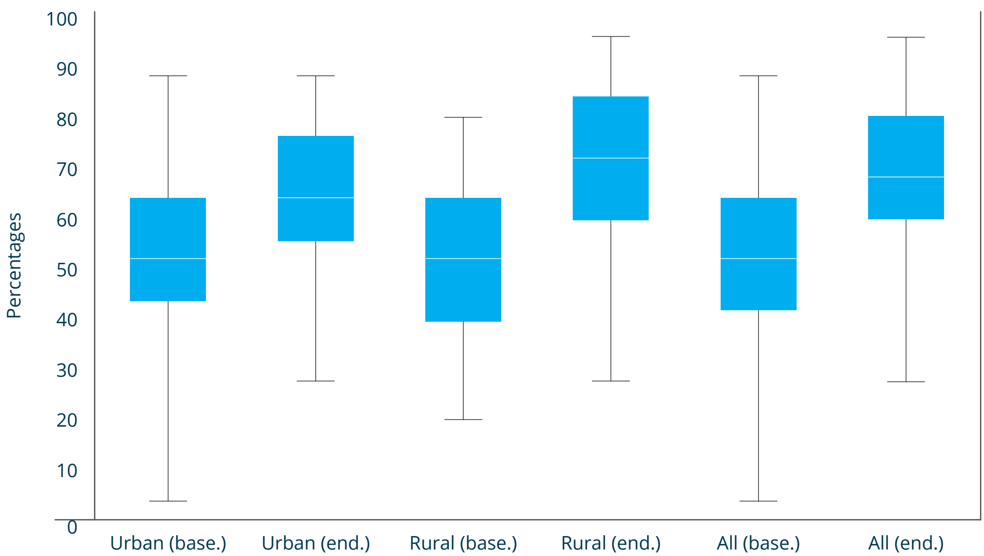
4.2 Impact on Grade R teachers

The impact on the Grade R teachers' PK, SMK and PCK was significant (paired samples t-test, $p < 0.001$), with a large effect size ($d = 1.37$). In the pre-test, the group mean was 51.4% and individuals obtained marks in the pre-test that ranged from 4.0% to 88.0%. In the post-test, the group mean was 68.1%, and individuals obtained marks ranging from 28.0% to 96.0%. See Figure 3.

The mean achievement in each district was similar in the pre-test: 50.7% (s.e. = 1.7) in the urban district and 52.4% (s.e. = 1.6) in the rural district. The difference in pre-test means of the two districts was not statistically significant (paired samples t-test, $p = 0.478$). In the post-test, the gap between the mean achievement in Urban District 1 (65.8%) and Rural District 2 (70.7%) widened. The increase of 15.1 percentage points in Urban District 1's mean achievement was statistically significant (paired samples t-test, $p < 0.001$) with a large effect size ($d=1.26$), as was the increase of 18.3 percentage points in Rural District 2's mean achievement (paired samples t-test, $p < 0.001$) with a large effect size ($d = 1.50$). However, the differences between the gain scores in Urban District 1 and Rural District 2 were not statistically significant (paired samples t-test, $p = 0.91$).

The wide range of performance in the test is understandable, given the disparate qualifications that the teachers had, which ranged from Matric only to postgraduate qualifications.

Figure 3 Grade R teacher performance in the pre-test and post-test by district and overall



4.3 Grade R teaching and support

The case-study findings indicated that teachers were using the R-Maths Programme materials[14] to plan and execute their lessons[15]. However, in some instances, the quality of lesson planning deteriorated, as teachers were relying heavily on the suggested daily lessons in the Activity Guides.

The Term Activity Guides – and the classroom kit, in particular – were highly appreciated and were being used extensively. Teachers appreciated that the resources and materials were available in the LoLT. The materials seemed to have increased teachers' enthusiasm to teach and learners' enthusiasm to learn mathematics.

The observed lessons were all relevant in terms of mathematics content areas and topics covered, mathematical concepts introduced, and skills practised. The lessons were largely aligned with CAPS and the R-Maths Programme's term plans outlined in the Activity Guides. The recommended lesson structure[16] was being loosely followed. Free-choice activities were often excluded due to lack of time. In several schools, teachers did not work with a focus group during the small group activities. Additionally, in some instances (more so at baseline), teachers did not clearly explain or model the group activities for the learners, and lessons ended abruptly without a conclusion.

There was some evidence that the application[17] of some of the R-Maths teaching principles[18] had improved (specifically, level and practice), but the application of the principles of level, guidance and interaction remained limited. The most applied principles were inclusivity and activity, followed by practice, whilst the least applied was level[19].

Most lessons were interactive and enjoyed by learners. Many of the teachers made effective use of questioning to engage the learners, but this was teacher-led. Learners rarely asked questions and in one lesson where they did, the teacher did not answer their questions straight away[20].

A positive finding was that support for Grade R teaching – both within the school and district – had increased. At baseline, limited support was provided to the Grade R teachers by the

14 These comprised: Training materials: A Facilitator's Guide and Teacher's Manual for each of the CTs and the block training; one Concept Guide (three sections: Overview of the R-Maths Project, Implementation of the R-Maths Project in the classroom and teaching the Five Mathematics Content Areas); four Term Activity Guides (one for each of Terms 1 to 4); one 'Big Poster Book' (with 11 double-spread posters); and one classroom kit.

15 The Maths focus time in the Grade R Daily Programme is 50 minutes.

16 The daily programme proposed by R-Maths comprises: Whole class activities (rhymes and songs, oral counting, counting concrete objects, and activities and questions linked to the Content Area focus of the lesson); small group activities (four independent groups and one teacher-guided group, with the groups rotating through five different group activities during the course of the week); and free choice activities (for enrichment activities).

17 Fieldworkers rated the extent to which the R-Maths principles were evident in the lessons which they observed using a five-point scale.

18 R-Maths advocates seven guiding principles of teaching Mathematics in Grade R which are: Context - learning takes place in meaningful and appropriate situations; Activity - learners should be directly involved in the learning-teaching process; Level - learners pass through various levels of understanding and development; Interaction - learning takes place when there is sharing and communication of ideas; Guidance - learning takes place when teachers guide learners in developing new knowledge; Inclusivity - learning takes place in an environment where everyone is welcomed, included, fairly treated, respected and can participate; and Practice - learning is consolidated through practising new skills and knowledge.

19 The ratings for the level principle increased from baseline to endline, but remained poor.

20 She returned to them later, when the learners no longer remembered what they had wanted to ask.

FP DHs. The scenario had improved somewhat at the endline: HODs reported attending an R-Maths onboarding workshop organised by the WCED, and HODs were reported to be providing increased support to Grade R teachers, including lesson observations.

Regarding district support, at baseline only one school reported that SAs conducted Grade R classroom visits. At endline, all schools reported that support from SAs included lesson observations (although their frequency was limited). Other forms of support from the district included: provision of a planning template, cluster workshops and WhatsApp groups that connected the SAs and Grade R teachers.

4.4 Impact on learners

The impact of R-Maths on the learners' mathematical knowledge and understanding was small but positive.

The tables below provide some summary statistics for the Marko-D test scores. The first is for the urban district and the second for the rural district.

Table 5 Marko-D descriptive statistics for the urban district

Group	Test phase	N	Minimum	Maximum	Mean	Standard error of the mean	Standard deviation
Intervention	Baseline	160	8.5%	87.2%	37.8%	1.2%	15.3%
	Endline	160	19.1%	95.7%	57.5%	1.3%	16.7%
Comparison	Baseline	157	2.1%	83.0%	35.0%	1.3%	16.5%
	Endline	157	14.9%	95.7%	55.3%	1.5%	18.9%

Table 6 Marko-D descriptive statistics for the rural district

Group	Test phase	N	Minimum	Maximum	Mean	Standard error of the mean	Standard deviation
Intervention	Baseline	150	2.1%	91.5%	42.0%	1.5%	17.9%
	Endline	150	12.8%	97.9%	58.4%	1.5%	18.6%
Comparison	Baseline	155	0.0%	89.4%	40.7%	1.3%	16.2%
	Endline	155	17.0%	87.2%	54.0%	1.3%	15.8%

In the urban district, the intervention group's baseline mean was 37.8% and increased by 19.6 percentage points (pp) to 57.5% at endline whilst the comparison group's baseline mean was 35.0% and increased by 20.3 pp to 55.3% at endline. The DiD of −0.7 pp indicates that the intervention group improved by 0.7 pp less than the comparison group. However, the net shifts were not significantly different for the two groups (independent samples t-test, $p = 0.371$).

In the rural district, the intervention group's baseline mean was 42.0% and increased by 16.3 pp to 58.4% at endline whilst the comparison group's baseline mean was 40.7% and increased by 13.3 pp to 54.0% at endline. The DiD of +3.0 pp indicates that the intervention group improved more than the comparison group. This change is statistically significant (independent samples t-test, $p = 0.035$) with a small effect size (d = 0.21).

Figures 4, 5 and 6 show box plots of the learners' performance in the pre-tests and post-tests, by group (for each district and then overall).

Figure 4 Grade R learner performance in the pre-test and post-test, by group (urban district)

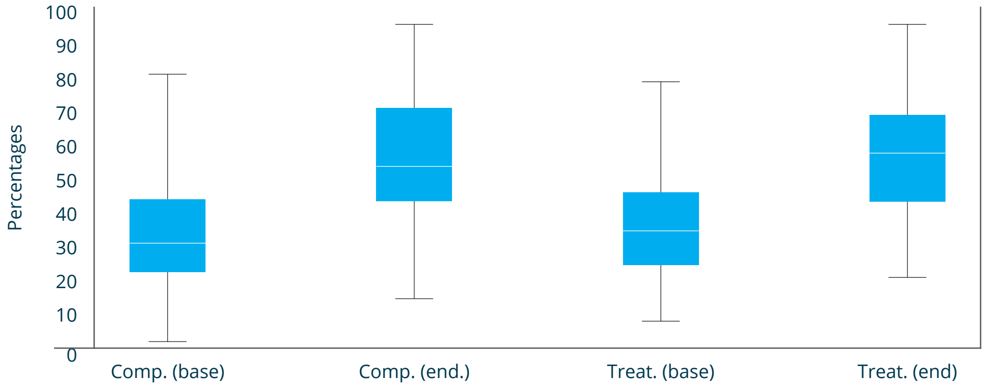


Figure 5 Grade R learner performance in the pre-test and post-test, by group (rural district)

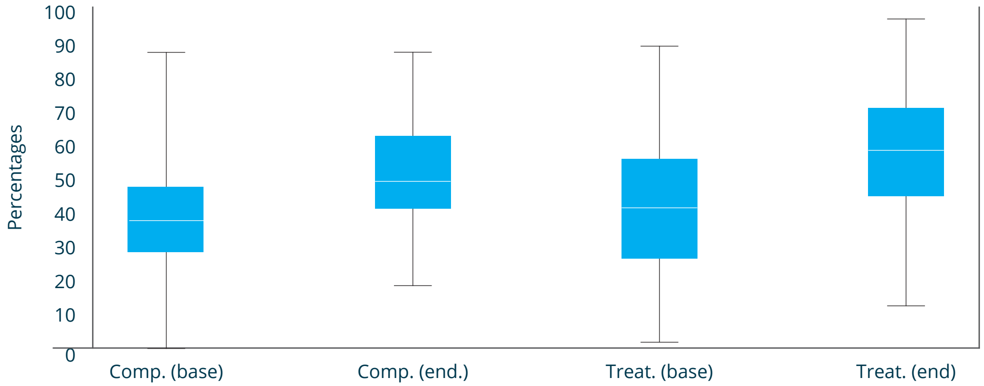
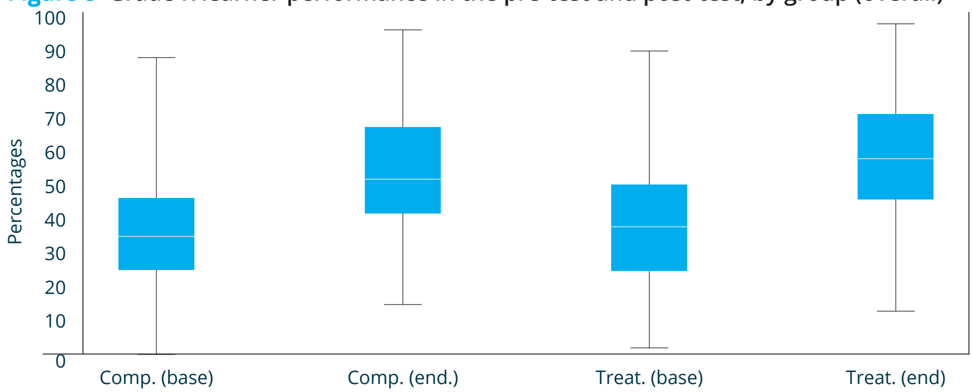


Figure 6 Grade R learner performance in the pre-test and post-test, by group (overall)



A GLM was conducted to control for differences such as age, gender, and other factors relevant to testing, to hold baseline scores constant and to explore differences between group, district, and language concurrently. The GLM found that the two factors that had the greatest effect on learner performance were district (medium effect, $d \approx 0.50$) and LoLT (medium effect, $d \approx 0.30$ to 0.50). Group (i.e. intervention or comparison) was also a significant factor, with the intervention group performing better than the comparison group: approximately 2.9 pp better over the whole Marko D test and approximately 5.0 pp better in the three levels of the test where the difference in improvement between the groups was significant. The relative improvement varied between 0.17 and 0.24 standard deviations. In all cases, the effect size was small ($d \approx 0.19$ to 0.24 depending on the test level). Older learners achieved higher average scores than younger learners at three levels of the test, but this difference was also small ($d \approx 0.20$). When comparing the average effect of the intervention and the average effect of age on Marko D performance, receiving the intervention was equivalent to about six additional months of age.

4.5 Embedding at province, district, and school levels

The consolidation phase had a specific focus on sustainability and the consolidation evaluation assessed the extent of embedding at province, district, and school levels.

The WCED was involved throughout in designing the R-Maths Programme and implementing R-Maths. Consequently, R-Maths expertise to provide training, support, and monitoring was developed. The WCED prioritised R-Maths and ensured that it was the only large-scale intervention implemented in the FP during 2017-2018. This meant that FP SA's could provide focused support and deliver the R-Maths training during this time. Ownership was evident in that R-Maths is integrated into the WCED's strategic plans and budgets and district improvement plans.

Initially, the FP SA's supported R-Maths implementation, but the initial evaluation highlighted that the SAs would not be able to continue to deliver training at scale or provide meaningful in-school support due to their workload and responsibilities[21]. Thus, lead teachers[22] were identified and upskilled to deliver training and provide in-school peer support. The evaluation highlighted concerns, however, with their selection and lack of incentives for lead teachers for these additional responsibilities.

R-Maths training was developed for and provided to FP HODs during the consolidation phase and, thereafter, FP HODs were required to submit a report on R-Maths implementation in their schools. A package of training, which includes literacy and R-Maths Programme materials, is now mandated for new FP HODs. R-Maths training was provided to new novice teachers during the consolidation phase and is now part of a three-year package of training for novice teachers entering the WCED. Through these activities there is now supplementary support for, and monitoring of R-Maths implementation at school level.

21 The consolidation evaluation report (Roberts & Mawoyo, 2020) estimates that with 75 FP SAs and an estimated 1000 primary schools, the WCED has a ratio of 13 primary schools to each FP SA.

22 'Lead Teachers' are teachers at schools who are selected to receive additional training for a particular Project to assist in supporting other teachers in aspects related to the Project (for example, lead subject meetings and planning workshops).

Fieldwork conducted for the consolidation evaluation in schools – three years after the start of implementation – found that R-Maths was well-understood, there was a high frequency and dosage of R-Maths implementation, and the activity guides and classroom kits were being used frequently. Assessment was taking place and remediation was implemented to address identified challenges. However, the concept guide was underused, and the R-Maths principles were not being consistently applied or explicitly referred to.

Thus, evidence was found that the R-Maths model had been tailored to the provincial and district context, embedded in planning and budgeting processes, and to a certain extent, institutionalised through integration into in-service training and professional development opportunities for existing and new teachers and FP DHs. The consolidation evaluation suggested that there was some distortion and dilution of R-Maths concepts and principles, which is unsurprising given that the training was being delivered via a third level in the cascade (lead teachers, trained by FP SAs, trained by the SDU).

5 Conclusion and the way forward

R-Maths sought to enhance the quality of Grade R mathematics teaching and learning in the Western Cape, by improving Grade R teachers' mathematics knowledge, understanding of the curriculum, and use of play-based pedagogies. It was rolled-out province-wide via a modified cascade model between 2016 and 2019, reaching 77 FP SAs and 3,800 Grade R teachers. SAs were trained first and subsequently provided training and support to Grade R teachers.

An impact evaluation found a positive impact on SAs' knowledge of supporting Grade R mathematics teaching, which increased on average by 17.5 percentage points from baseline to endline in a customised test. The finding was statistically significant with a large effect size. However, the SAs' endline test scores remained varied, meaning that support for Grade R mathematics teaching is likely to be uneven. A positive impact was also found on Grade R teachers' knowledge of Grade R mathematics teaching, which increased similarly, by 17.4 percentage points from baseline to endline. The result was statistically significant with a large effect size. The endline test scores were very wide-ranging, which is likely to manifest in a highly varied quality of Grade R mathematics teaching.

A small but positive impact was found on the Grade R learners' mathematical knowledge and understanding. A GLM constructed to control for confounding variables found that learners whose teachers had participated in R-maths performed at endline 2.9 percentage points better than learners whose teachers had not. This finding was statistically significant with a small effect size.

Support for Grade R teaching had increased, both within schools and to schools by education districts, following R-Maths implementation.

Some evidence was found that the roll-out of R-Maths in the Western Cape was successful in improving teaching and learning. Thus, the R-Maths Programme materials, delivered via a modified cascade model (which built the capacity of SAs to train and support Grade R teachers), holds promise of being a model that works at scale. R-Maths was subsequently rolled out to all Grade R teachers in the province and has been embedded as a core initiative of the WCED.

What recommendations for the design and roll-out of other large-scale education interventions in South Africa can be gleaned from what was learned through the implementation and evaluation of this project? Key learnings – based on both the data analysis and the experiences of the implementation and evaluation teams – are outlined below. They have been written as broad principles (to be relevant to almost any large-scale school-level education intervention in South Africa).

1. Province-wide interventions should be located and managed within the provincial Curriculum Directorate. This will facilitate buy-in, ownership, and sustainability once the funders and external service providers exit.
2. A Project Steering Committee, comprising senior representatives of the provincial education department and key project partner organisations, should be established during the planning stage of the project, with adequate attention to building a good level of trust between all role-players. This will ensure that all stakeholders are involved in decision-making. The Project Steering Committee should ideally be chaired by a high-ranking education department official.
3. The involvement of an expert external support team – the role played by the SDU in the Western Cape – is essential. This team will have the knowledge, skills, and time needed to support implementation by over-stretched departmental officials.
4. Upskilling departmental officials in subject-content knowledge and involving them in project roll-out and training is one of the best ways of ensuring project sustainability over the medium-term to long-term.
5. Roll-out requires sufficient highly motivated departmental officials that are not too overloaded with other tasks and have adequate knowledge and skills.
6. The number of schools allocated to each department official to support should be not too high, as this would preclude them from providing adequate support to their allocated schools.
7. Department officials do not have the capacity to monitor and support implementation in the classroom to any great extent and therefore school-based support for the implementation of what has been learned in day-to-day teaching is necessary. This is ideally provided by in-situ suitably-trained individuals, like FP HODs for whom this is a part of their core responsibilities.
8. Lead teachers – carefully chosen for their high levels of competence and excellent performance in the pre-project assessment – should play a key role in the roll-out.
9. To ensure sustainability, newly appointed education officials, DHs and teachers should be provided with the same training as the original cohort of people trained. In practical terms this means planning and budgeting for training to be offered in subsequent years.
10. Only one province-wide intervention should be permitted in a phase of schooling at a time, both during implementation and for some time (at least one year) thereafter.
11. Interventions should be evaluated rigorously to add to the available evidence base to determine which programmes are effective in improving outcomes. This will ensure that ineffective and non-cost-effective programmes are not replicated.
12. Where it is feasible and will not hinder the intervention, interventions should be designed so that it is possible to obtain randomly assigned treatment and control groups.

Acknowledgements

1. The Zenex Foundation, the ELMA Foundation, and the Maitri Trust, for funding the evaluations.
2. Matthew Snelling, for analysis of the learner data.
3. Nicky Roberts and Monica Mawoyo, for work on the consolidation evaluation
4. Stephen Taylor, for reviewing and suggesting how to improve this chapter.
5. Hogrefe and the University of Johannesburg, for permission to use the Marko D learner test in the evaluation and translate it into isiXhosa.

References

Clements, D., & Sarama, J. (2014). *Learning and Teaching Early Math: The Learning Trajectories Approach*. London: Routledge.

Cross, C. T., Woods, T. A., & Schweingruber, H. E. (2009). *Mathematics learning in early childhood: Paths toward excellence and equity*. Washington, DC: National Research Council.

Department of Basic Education. (2011). *Curriculum and Assessment Policy Statement (CAPS). Mathematics*. Pretoria: Department of Basic Education.

Department of Basic Education. (2015). *The South African National Curriculum Framework for Children from Birth to Four.* Pretoria: Department of Basic Education.

Fleisch, B. (2018). *The education triple cocktail: System-wide instructional reform in South Africa*. Cape Town: UCT Press.

Fleisch, B. & Motilal, G.B. (2020). The triple cocktail programme to improve the teaching of reading: Types of engagement. *South African Journal of Childhood Education*, 10(1), 1–13.

Fritz, A., Balzer, L., Ehlert, A., Herholdt, R., & Ragpot, L. (2014). A mathematics competence test for Grade 1 children migrates from Germany to South Africa. *South African Journal of Childhood Education*, 4(2), 114–133.

Fritz-Stratmann, A., Ehlert, A., & Klüsener, G. (2014). Learning support pedagogy for children who struggle to develop the concepts underlying the operations of addition and subtraction of numbers: the 'Calculia' programme, *South African Journal of Childhood Education*, 4(3), 136–158

Gove, A., Korda Poole, M., & Piper, B. (2017). Designing for Scale: Reflections on Rolling Out Reading Improvement in Kenya and Liberia. *New Directions for Child & Adolescent Development*, 2017(155), 77–95.

Hazell, E., Spencer-Smith, G., & Roberts, N. (2019). Improving Grade R mathematics teaching in South Africa: Evidence from an impact evaluation of a province-wide intervention. (University of KwaZulu-Natal), 76, 48–68.

Jensen, H., Pyle, A., Zosh, J. M., Ebrahim, H. B., Scherman, A. Z., Reunamo, J., & Hamre, B. K. (2019). *Play facilitation: the science behind the art of engaging young children* (white paper). Billund, Denmark: The LEGO Foundation.

Kotze, J., Fleisch, B., & Taylor, S. (2019). Alternative forms of early grade instructional coaching: Emerging evidence from field experiments in South Africa. *International Journal of Educational Development*, 66, 203–213.

Kuhne, C., Lombard, A.P., & Moodley, T. (2013). A learning pathway for whole numbers that informs mathematics teaching in the early years. *South African Journal of Childhood Education* 3(2), 77–95.

Kuhne, C., van den Heuvel-Panhuizen, M., & Ensor, P. (2005). Learning and Teaching Early Number: Teachers' Perceptions. In Chick, H. L. & Vincent, J. L. (Eds.). *Proceedings of the 29th Conference of the International Group for the Psychology of Mathematics Education* (PME), 3, 217–224.

McDevitt, D. (1998). How Effective is the Cascade Model as a Method for Disseminating Ideas? A Case Study in Botswana. *International Journal of Educational Development*, 18(5), 425–428.

Montie, J. E., Xiang, Z., & Schweinhart, L. J. (2006). Preschool experience in 10 countries: Cognitive and language performance at age 7. *Early Childhood Research Quarterly*, 21, 313–331.

Pyle, A., DeLuca, C., & Danniels, E. (2017). A scoping review of research on play-based pedagogies in kindergarten education. *Review of Education*, 5(3), 311–351

Roberts, N., & Mawoyo, M. (2020). R-Maths consolidation phase evaluation report. (Unpublished). Johannesburg: Kelello.

Sarama, J., & Clements, D.H. (2009). *Early Childhood Mathematics Education Research: Learning Trajectories for Young Children*. London: Routledge.

Shonkoff, J. P., Richmond, J., Levitt, P., Bunge, S. A., Cameron, J. L., Duncan, G. J., & Nelson III, C. A. (2016). *From best practices to breakthrough impacts: A science-based approach to building a more promising future for young children and families*. Cambridge, MA: Harvard University, Center on the Developing Child, 747–756.

Suzuki, T. (2008). *The Effectiveness of the Cascade Model for In-service Teacher Training in Nepal*. Available at: https://www.semanticscholar.org/paper/The-effectiveness-of-the-cascade-model-for-teacher-Suzuki/80b0aa96e35f2e8dd175969c0f1a047144381101?p2df (Accessed 14 November 2022).

Sylva, K., Taggart, B., Siraj-Blatchford, I., Totsika, V., Ereky-Steven, K., Gilden, R., & Bell, D. (2007). Curricular quality and day-to-day learning activities in pre-school. *International Journal of Early Years Education*, 15(1), 49–65.

Sztajn, P., & Wilson, P.H. (2019). *Learning Trajectories for Teachers: Designing Effective Professional Development for Math Instruction*. New York: Teachers College Press.

UNICEF. (2020). Glaser, M. & Biersteker, L. (Eds). Play: Teacher Training Packages. NQF levels 4, 5, 6 and 7. Facilitator and Student Guides. Play Learning Materials Consortium (PLMC). (Unpublished). Pretoria: UNICEF.

Van den Heuvel-Panhuizen, M., Kuhne, C., & Lombard, A.P. (2012). *The Learning Pathway for Number in the Early Primary Grades*. Midrand: MacMillan.

Wedell, M. (2005). Cascading training down into the classroom: The need for parallel planning, *International Journal of Educational Development*, 25(6), 637–651.

White, H., & Sabarwal, S. (2014). Quasi-experimental Design and Methods, *Methodological Briefs: Impact Evaluation* 8. Florence: UNICEF

Whitebread, D., Kuvalja, M., & O'Connor, A. (2015). *Quality in Early Childhood Education: An International Review and Guide for Policy Makers.*

Whitebread, D., Neale, D., Jensen, H., Liu, C., Solis, S.L., Hopkins, E., Hirsh-Pasek, K., & Zosh, J. M. (2017). *The role of play in children's development: a review of the evidence (research summary)*. Billund, Denmark: The LEGO Foundation.

Zosh, J.M., Hirsh-Pasek, K., Hopkins, E.J., Jensen, H., Liu, C., Neale, D., Solis, S. L., & Whitebread, D. (2018). Accessing the inaccessible: Redefining play as a spectrum. *Frontiers in Psychology*, 9: 1124.

Zosh, J. N., Hopkins, E. J., Jensen, H., Liu, C., Neale, D., Hirsh-Pasek, K., Solis, S. L., & Whitebread, D. (2017). *Learning through play: a review of the evidence*. Billund, Denmark: LEGO Foundation

11

The implementation of the Bala Wande programme in Grade 1 in three provinces: Lessons learned

INGRID SAPIRE, PERMIE ISAAC, SIHLOBOSENKOSI MPOFU, LORNA SAKO, MMAPULA SEOLOANA, THOBEKA NDAMASE & VUYOKAZI MAFILIKA

Abstract

In this chapter we use excerpts from the Bala Wande (BW) materials to give insight into the planning, structure, and affordances of the programme print material, present some representative causal results in Grade 1 classes from two of the implementation sites, and report on lessons learned from the take-up of the Bala Wande mathematics programme based on classroom observation of 19 high-agency teachers in Grade 1 schools across three implementation sites. The programme, designed to take into account the limitations of scripted lessons, provides print materials in the form of a teacher guide (TG) and a learner activity book (LAB), mathematical manipulatives, and teacher support (various support modalities including training, videos, teacher assistants, and coaches). The evidence presented in this chapter shows that there is take-up of the programme materials, even in the early stages of implementation. Quantitative findings show evidence of the programme working, albeit in small increments, in what is still 'early days' and in spite of Covid-19. Classroom observations show that high-agency teachers in all provinces demonstrated a relatively high degree of fidelity to the programme materials. Lessons learned from the take-up speak to materials design and

KEYWORDS
Early grade maths, teacher support, materials design, programme implementation, structured pedagogy

implementation support. Take-up findings thus begin to address the question of *What can be scaled?* although more time is needed to gauge the full effect and consider all of the implications for programme review.

1 Introduction

The South African education system has been described as one in crisis (Fleisch et al. 2017; Spaull 2013; Howie & Chamberlain 2017). Earlier interventions that aimed to address the crisis focused on solving issues that related to inequalities (for example equalising school facilities, class size, and materials issued) but more recently the focus has shifted to instruction, with much emphasis on the relationship between teachers, students, and the resources they use. Amnesty International (2020), in their report "Broken and Unequal", expose many challenges faced by the education system in South Africa. Among these is the challenge that teachers lack the requisite knowledge and skills in subjects they teach, exposing a clear need for intervention in early grade mathematics (EGM). Early interventions have distinct advantages – particularly for students with difficulties in learning (Gersten et al. 2009). Many authors point to the need for systemic EGM interventions because studies show that students who struggle in the early grades find it difficult to perform better in the higher grades (Aunola et al. 2004; Morgan et al. 2009).

2 Addressing challenges in mathematics education in the early grades

2.1 Literature review

The challenges present in EGM are not unique to South Africa – there are general challenges that students face in EGM linked to counting, number knowledge, number transformation, and estimation (Jordan et al. 2006). While some of the literature speaks to student-directed interventions, the Bala Wande (BW) intervention targets teachers, providing the material as described above as well as other support which will be discussed in more detail later in this chapter. Research has shown that some teachers struggle with both pedagogy and subject knowledge (Christiansen et al. 2018) hence consideration of these needs is critical.

Some of the challenges in relation to mathematics learning are attributed to how students are taught rather than the difficulty of the mathematics itself. Factors that contribute to learning difficulties include the type of teacher-student interactions, use of mathematical language, and teaching method (Brown et al. 1998). Teacher professional development has been identified as one of the ways that can be used to address the challenges of poor results in early mathematics (Sowa et al. 2021). The focus of the mathematics interventions should be on 'big ideas' such as meaningful counting and a focus on the structure of numbers (Clements & Sarama 2014).

Teacher talk plays a critical role in the success of students' learning of mathematics. Resnick et al. (2015) advocate for classroom talk that develops students' ability to think, instead of students just listening to the teacher and giving correct answers to answer-driven questions. Students who are exposed to a more accountable classroom talk have shown better gains in the subject of study than those who are exposed to answer-directed talk (Resnick et al. 2015). Attention needs to shift from 'do classrooms have student talk?' to 'what kind of talk do teachers elicit from their students?' as there is evidence that the type of student talk determines the type of learning happening in the classroom. Mehaan has classified another type of teacher-dominant talk which follows a typical pattern: teacher initiation, student response, and teacher evaluation (Meehan 1979). Although this type of activity involves all students, it is usually through a choral answer, and many students are physically present, passively engaged, but psychologically absent (Hattie 2012). Another important element of teacher talk relates to questioning. Ellis (1993) claims that many teachers rely on low-level cognitive questions in order to avoid a slow-paced lesson, keep the attention of the students, and maintain control of the classroom. Probing questions classified as high-level cognitive questions (Ellis 1993) have the potential to reveal whether a student has grasped a concept as they call on deeper understanding of the topic (Brualdi 1998).

Independent work is another aspect of classroom activity that is necessary in support of meaningful learning. Student writing in mathematics does not only document what has been happening in the classroom, it also provides opportunities to deepen students' understanding and has potential to give a new perspective to students in the skills and concepts that the teacher seeks to develop (Urquhart 2013). Although it does happen in some classrooms, Alvermann (2002) suggests that it is unimaginable to teach students without having them write in mathematics. Alvermann goes on to state that writing raises the cognitive bar and improves thinking and reasoning skills. Pugalee (2005) suggests that writing helps students develop problem-solving skills and supports both mathematical reasoning and communication.

2.2 On the ground: Bala Wande – an EGM intervention programme

Bala Wande (BW) is an intervention programme designed to investigate ways in which materials can be used to support the teaching and learning of mathematics. The programme trials the use of materials using different support modalities to provide evidence of different approaches to improving EGM outcomes in South Africa. From the outset the programme chose to take a collaborative-development approach: print material and videos were developed in conversation with a reference team made up of representatives from tertiary institutions, NGOs, teachers, district officials and the Department of Basic Education (DBE). The programme is CAPS-compliant in terms of content, time allocation and assessment although the materials follow a reorganised curriculum which has clustered the content topics[1]. For example the topic of mass is taught in one full week (clustered curriculum design) rather than being spread out over the four terms of the year (spiral curriculum design). The development of the materials also drew on relevant mathematics education literature. Observations from implementation in the field have been used to inform the ongoing review and refinement of the materials which are currently in pilot form.

The print material consists of a teacher guide (TG), which provides day-by-day guidance for teaching of mathematical concepts, and a learner activity book (LAB) with worksheets and games to be used for consolidation of concepts that have been taught in each lesson. There are also manipulatives boxes for teachers and students. Teachers receive a box with a set of posters and demonstration-size manipulatives (e.g. large number-symbol and number-name flash cards, a large magnetic ten-frame with magnetic counters, a demo bead string, etc.). Student boxes contain the same manipulatives (student size) with sufficient for each student to have and hold their own manipulatives when doing the lesson activities. The manipulatives are scripted into the lessons so that teachers will know when and how to use them and be sure to involve students in using them as well. Teachers also receive a bilingual dictionary of mathematical vocabulary and there are videos of classroom footage that bring to life the photo-scripted lesson activities to assist teachers when they prepare. Some of the features of the material are now discussed and exemplified.

In the BW materials, conceptual learning is supported through a methodology that is embedded in the planned lessons. The TG methodology is presented using photo sequences, which give visual prompts about what is needed when teaching the daily lessons and thus to cut back on reading. The activities are carefully sequenced and provide guidance for the scaffolding of conceptual teaching. For example, in the first weeks of teaching the foundation is laid for using 1-1 correspondence when counting objects to establish number concept. Figure 1 shows a photo-scripted lesson activity which centres around a farm-scene poster that is also used for consolidation that day in the LAB.

1 In a similar manner to the TMU (Teaching Maths for Understanding), a pilot implementation of the Mathematics Framework (DBE 2018) follows a reorganised CAPS which clusters certain topics to allow more extended time per topic. This is also in line with the curriculum design of many other African and international mathematical curricula.

Figure 1 **Excerpt from the Bala Wande isiXhosa Teacher Guide (Grade 1, Term 1, pages 43-44)**

Source: Bala Wande Afrikaans Learner Activity Book Grade 1 Term 2, page 6. From: https://fundawande.org/learning-resources?i=8#learningResourcesHolder

The photo-script in Figure 1 shows the teacher using the poster, magnetic ten-frame and counters to find out how many of each of the farm animals appear on the poster. The script shows that the teacher should call students to the board to place counters over the named animal to find out 'how many' of each one there are. The script is minimal as the visual prompts in the photographs support the script.[2] The script leads into the independent work where students work in their LABs on a similar activity, using their own counters. The 'hand' inserted on the learner page (with a red border) is a further prompt (language free) to show students to match counters to animals in order to find out 'how many' of each one there are.

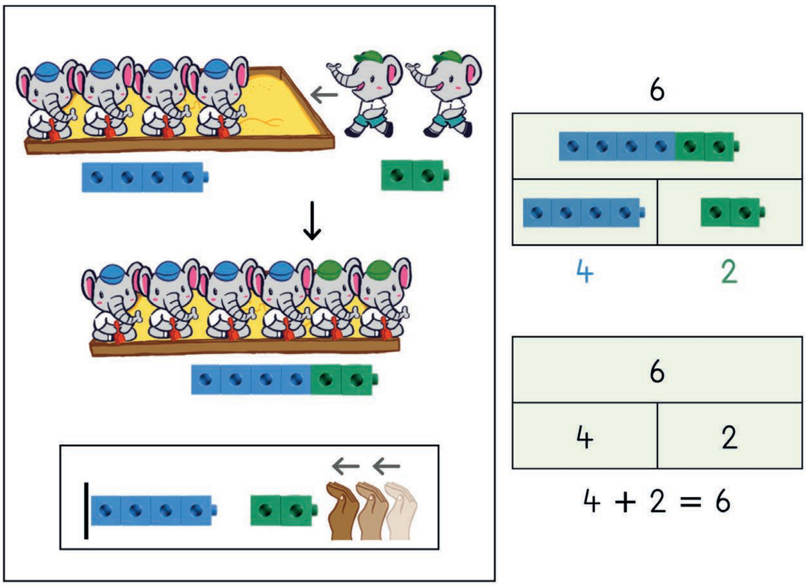
The BW materials introduce numbers and symbols at an early stage, using concrete objects or their representations. For example, there are lessons that are dedicated to writing number symbols which simultaneously develop students' number sense since learners are engaging with counts and number symbols at the same time. A lesson 'script' for this type of lesson activity will show the teacher engaging actively with the class, using manipulatives such as multifix blocks or bead strings, writing on the board and encouraging students to do activities contained in the aligned activities in the LAB.

The LAB promotes students' written work – it provides worked examples in support of teacher's explanations, mediation, and students' independent work. Mathematical representations help illuminate mathematical concepts that may be

2 The BW videos offer further support – a short clip of a teacher working through the activity with a class.

difficult to comprehend in abstract form. Representations also help students organise their mathematical thinking and support abstraction of concepts. The BW materials developers applied this theory, and the LAB is richly illustrated to support the move to abstraction. Figure 2 shows the way in which the introduction to addition is represented in one instance in the LAB.

Figure 2 **Excerpt from the Bala Wande Learner Activity Book (Grade 1, Term 2, page 6)**

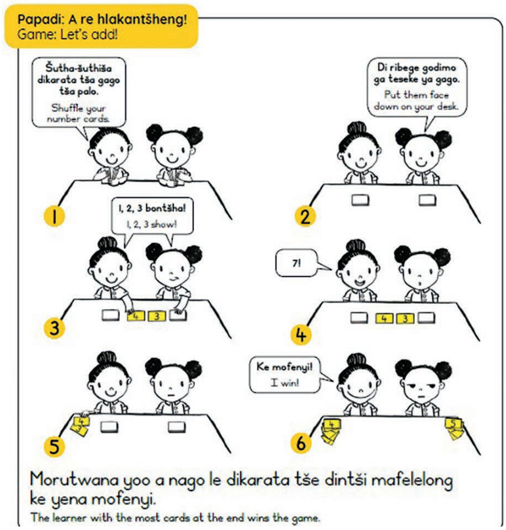


Source: Bala Wande Maths Learner Activity Book (Grade 1 Term 1, page 6.) From: https://fundawande.org/learning-resources?i=8#learningResourcesHolder

This graphic illustration provides multiple representations of 'adding on'. The elephants, one of the 'BW creatures' are shown with blue and green caps, to align with the blue and green multifix blocks. The hands show that two green blocks are being added on to the four blue blocks. The blue and green blocks in the number table are linked with blue and green symbols before the more abstract representation of numerals in the second number table is given. In the scripted lesson that introduces this independent activity, the teacher uses blocks and number tables drawn on the board.

Active learning is promoted in the BW materials, not only through the promotion of interaction between teachers and students in every mathematics lesson, but also through the inclusion of games, many of which are played with manipulatives that help students to visualise and explore numbers and other mathematical concepts in a variety of contexts. The Grade 1 games are exciting and varied and they provide opportunities for learning and consolidating of the various concepts. Figure 3 shows a game where the number cards are used to add.

Figure 3 Excerpt from Bala Wande Sepedi Maths Learner Activity Book (Grade 1 Term 2, page 27)



Source: Bala Wande Maths Sepedi Learner Activity Book (Grade 1, Term 2, page 27) From: https://fundawande.org/learning-resources?i=8#learningResourcesHolder

The speech bubbles in the game cartoon are included to encourage interaction while the illustrations support the instructions of how to play the game. In this game, students shuffle their cards, then they each put one down. The one who is first to give the sum of the numbers shown, keeps the cards.

One of the key affordances of the BW photo-scripted TG is that it supports verbal interaction in the classroom. The speech bubbles exemplify both the instructions to be given and questions teachers can ask when engaging with students. They also demonstrate ways in which students could respond. This is designed to create an awareness of conversations that could take place in a class in line with best practice suggested in the literature. The limitation of space is felt strongly in the development of the scripts. This is one aspect of the TG that will be carefully reviewed based on feedback from the implementation sites.

3 Data

In this chapter we report on two sets of data. The first is from the programme-external evaluation and is presented to shed light on effects of programme implementation in the first year. The second is video observation data collected to show evidence of programme take-up by high-agency teachers. Analysis and findings of the evaluation data is presented first, followed by that of the video observation data. Further detail on each data set is included below to ensure coherence of presentation.

4 Analysis and findings

4.1 Evidence of programme effects in the first year

The implementation of the BW intervention commenced at three sites in January 2021, the Eastern Cape (EC), Limpopo (LP), and the Western Cape (WC). Effects of Covid-19 were still felt at this time and learners were still attending school according to a rotation system. The external evaluation team collected baseline data in the EC in the first term of 2021 and in LP schools in the third term of 2021 (the delay in LP was due to Covid impacting on school attendance in the province).[3] The external evaluation reports based on this data (LP and EC) are used in this chapter to address our first research question.

The impact evaluation in LP uses a randomised control trial (RCT) with 120 no-fee schools in Capricorn North and Capricorn South districts randomised into one of three arms. The three arms of the LP intervention (40 schools per arm) are a control group, a group that receives materials only (LTSM), and a group that receives materials and also has TAs in each class (LTSM+TA). The LTSM and LTSM+TA arms also receive training and are monitored by the programme team. There are around 2,400 Grade 1 learners across the 120 schools. In the EC there are 29 intervention schools and 30 control schools. Within each school, 16 Grade 1 learners were randomly selected to form the EC cohort. Two different instruments were used to assess mathematics skills. The first was a 1-1 numeracy assessment; the second was a written assessment administered to groups of up to ten learners at a time.[4] Figures 4 and 5 present results from LP and Figure 6 presents results from the EC, showing evidence of early impact of the programme.

3 Data was not collected in the WC.

4 This instrument was developed for Grade 1 learners by Brombacher and Associates and designed to be administered to groups of learners using tablets.

Figure 4 LP estimated effect sizes for early numeracy (standard deviation) (baseline)

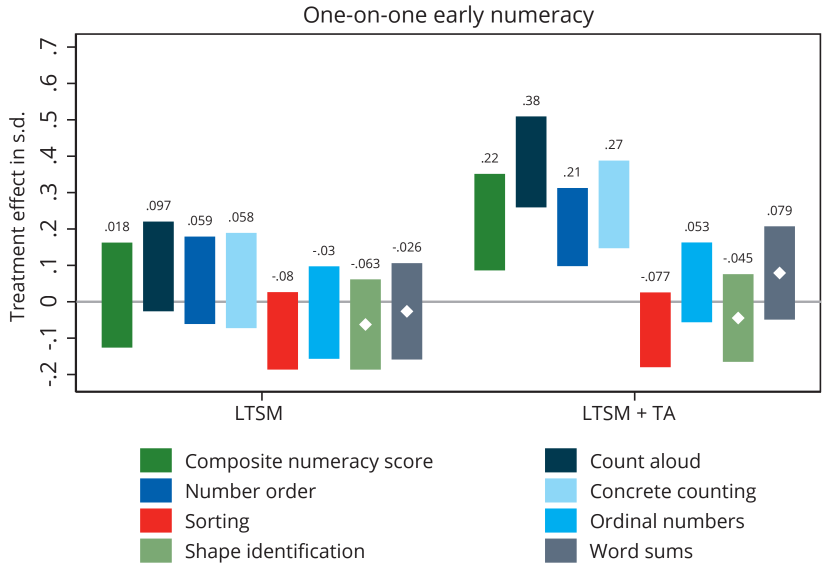


Figure 4 shows that the overall effect size is 0.22 standard deviations. This is evidently driven by significantly better performance on the procedural numeracy tasks of counting, number identification, and ordering. We see no difference in performance on the more conceptual tasks such as sorting, ordinal numbers, shapes, or word sums.

Figure 5 LP estimated effect sizes for EGMA (standard deviation) (baseline)

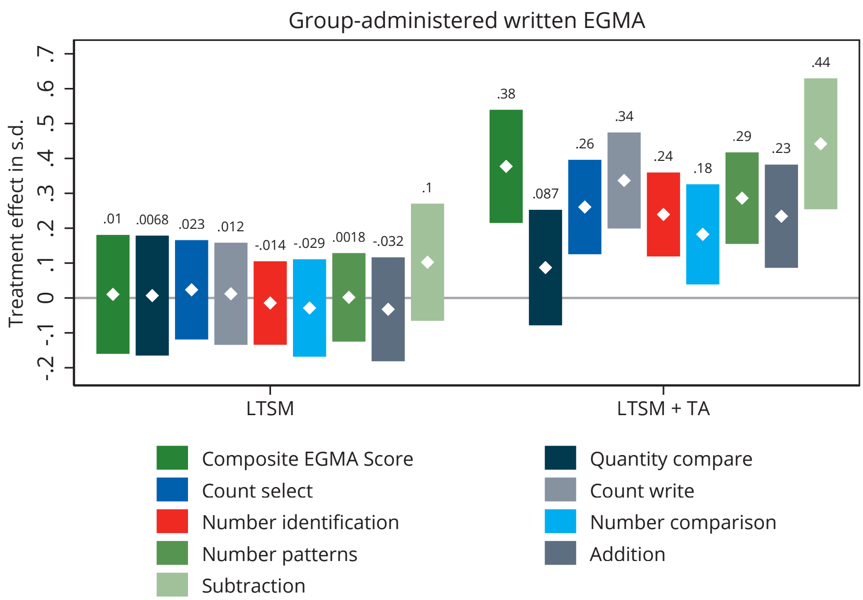


Figure 5 shows the estimated effect sizes for composite EGMA score and the standardised version of the component tasks. Learners in LTSM+TA schools do not score significantly higher than control learners on the quantity comparison task. On all other tasks, they outperform their peers in control schools, with effect sizes ranging from 0.18 standard deviation to 0.44 standard deviations. The effect size on the composite EGMA measure is 0.38 standard deviations.

Figure 6 EC estimated effect sizes for early numeracy (standard deviation - midline reporting)

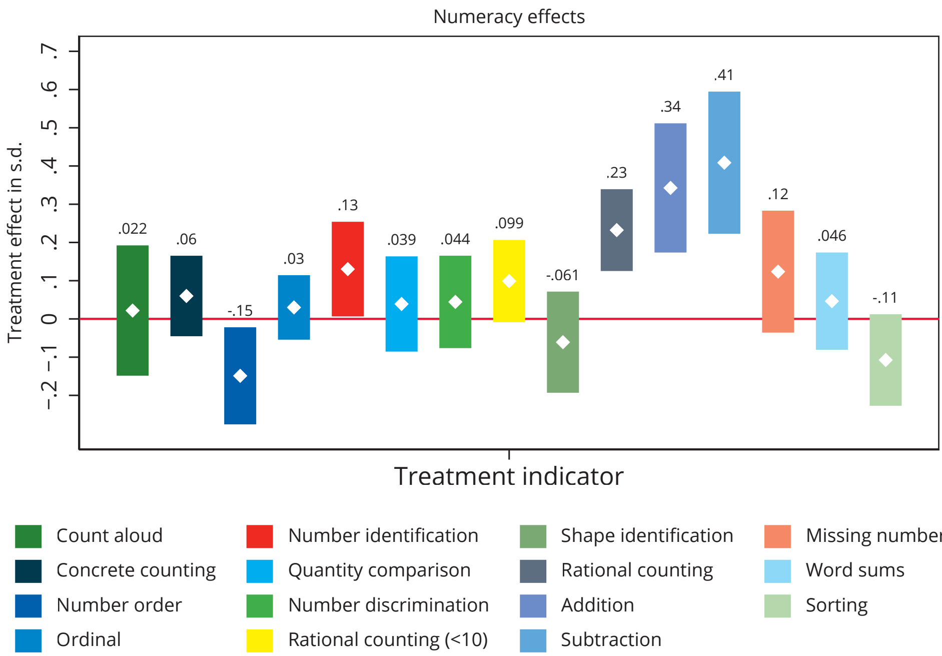


Figure 6 shows mixed evidence of the impact of BW on basic number concept using baseline-midline data. Most of the subtasks have small and insignificant effects, except for two subtasks where the difference between treatment and control is marginally statistically significantly different (at the 10 per cent level). The first is a negative effect of 0.15 standard deviations on number ordering and the second a positive effect of 0.13 standard deviations on number identification. There is a highly statistically significant (at the 1 per cent level) positive impact of 0.24 standard deviations on the more challenging rational counting task. This final task lies somewhere between basic number concept and procedural fluency as use of efficient methods relies on some strategic competence. The BW intervention has substantial positive effects of 0.35 and 0.42 standard deviations on the procedural fluency tasks of addition and subtraction respectively.

Taking into account the findings shown above, positive effects of the programme are being felt, which need to be consolidated over time. What has been shown to date is that the BW schools significantly outperformed their peers in control schools on subtasks that assessed procedural fluency. We also found a significant positive impact on the rational counting of larger quantities, a skill in the basic number concept domain but with reliance on some strategic competence.

4.2 Evidence of take-up amongst high-agency teachers in programme schools

The BW maths specialist team collected video data in order to learn about programme intervention from a 'best practice' sample of high-agency teachers. Our methodology was design-based research (DBR) since in carrying out this research we considered ways in which the BW programme, the 'intervention', works as a proposed 'solution' to the problem of low outcomes in EGM (Collins 1990; Brown 1992). We applied a mixed-methods approach since we used both quantitative and qualitative data analysis to inform our findings. Molina-Azorín & Cameron (2010) state the mixed-methods approach might provide better understanding of the research questions. The following research questions guided our data collection, analysis and discussion:

- How do high-agency teachers use the BW materials?
- What lessons can be learned from the take-up of high-agency teachers?

The sample was made up of video observations of 19 high-agency teachers identified across the sites by the programme managers. All the necessary consent was obtained, at all levels. Observation data was video recorded because we intended to "reflect on teacher practices for further interpretation and analysis" (Griffin 2017). Video data also offers greater transparency (Heath et al. 2010). We collected five videos from the EC (240 minutes), seven videos from the WC (325 minutes) and seven videos from LP (380 minutes). Video data records more information than field notes or audio recordings and it facilitates analysis at a later stage (Marra & Holmes 2008).

The videos were coded using five-minute time segments taken across the entire time of videoed data for each observed lesson (195 segments) for the purpose of quantitative analysis. The unit of analysis was thus a five-minute time segment. The length of time segment was chosen in order to give fair representation to activities occurring in the lesson, related to each of the 29 variables being coded.[5] To ensure validity, initial coding was done by the whole team participating in consensus discussions during which codes were refined and agreed on. After this, to ensure reliability, at least three team members were involved in the coding of the remaining videos and one team member was present at all coding discussions. After cleaning the data we drew up tables and graphs which we used to identify core themes to inform the qualitative discussion of the data.

We present findings and discussion on a subset of the coded variables that are relevant to take-up of the programme in the early stages of implementation across the three province sites.

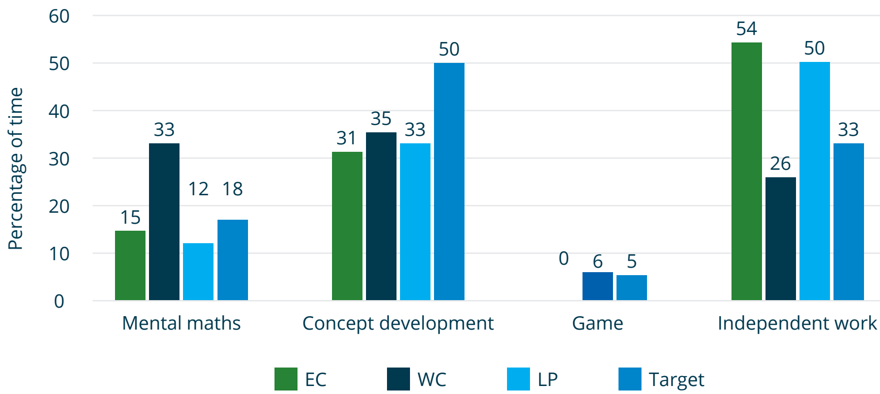
4.2.1 BW activity type

The TG suggests approximate timing for the three main scripted activities. These are guidelines to suggest an appropriate balance of time on different activities, but it is not expected that teachers follow the guidelines exactly. The programme encourages teachers to use their professional discretion and make choices in the best interests of

5 The full list of operationalised codes cannot be included here due to space constraints.

teaching and learning. The suggested 'target' times are: mental maths (15 minutes – 17%), concept development (45 minutes – 50%) and independent work (30 minutes – 33%).[6] It can be seen in Figure 7 that there were similarities and differences in the amount of time spent on the scripted activities across the sites.

Figure 7 **Bala Wande activity type**



In LP and EC time spent on mental maths lesson-starter activities is close to the expected time (12% and 15% respectively) while in the WC relatively more time is being spent on mental maths (33%).[7] A typical mental maths lesson begins with a register, calendar, and weather discussion, followed by an activity or game as illustrated in Figure 8.

Figure 8 **Excerpt from the Bala Wande Grade 1 Teacher Guide (Term 2, page 156) (on left) and a screenshot of a teacher conducting the same activity in her classroom (on right)**

Source: Bala Wande Grade 1 isiXhosa Teacher Guide Term 2, page 156. From: https://fundawande.org/learning-resources?i=8#learningResourcesHolder

6 These percentages are shown as 'target' percentages in Figure 7 though as suggested teacher may not stick exactly to these times, this may vary from day to day and between teachers. Games were also coded here (activity type) although they are not scripted separately. They are included in concept development, independent work and mental maths activities and we were interested in seeing how much time was spent playing games.

7 Here and in all reporting on video coding, percentages express percentages of total time observed.

On further investigation of the type of mental maths activities that were being done it could be seen that some teachers did more than one activity, usually by combining a few of the scripted mental maths activities. Another variation was that rote counting to 100 and skip counting in 2s, 5s, and 10s at the beginning of a lesson was also included (more so in the WC).

Lessons learned:

- The reiteration of the explanation of the purpose of the BW mental maths activities combined with watching videos of the activities and role-playing them during training (LP and EC) has had a positive impact on the type of mental maths activities being done as well as the duration thereof.
- Training sessions should explicitly highlight the importance of adhering to time and navigating the TG to see which activity has been scripted for that day.
- Further thought could be given to supporting teachers who have minimal training (such as in the WC) – platforms such as WhatsApp could be used.

In the concept-development activities, the teacher introduces a new concept or skill that will be consolidated in independent work. Figure 7 shows that time spent on concept development was similar across the three sites (EC: 31%; LP: 33%; WC: 35%), all below the 50% suggested in the scripted plans. There was evidence that the TG supported quality interactions between teachers and learners, an example of this is shown in Figure 9.

Figure 9 **Excerpt from the Bala Wande Teacher Guide (Term 2, page 181) (on left) and a screenshot of a teacher conducting the same activity in her classroom (on right)**

Source: Bala Wande Grade 1 isiXhosa Teacher Guide Term 2, page 181. From: https://fundawande.org/learning-resources?i=8#learningResourcesHolder

According to our observation concept development is given insufficient time. The reduction of this time deprives learners of the scaffolding that is aimed at building understanding of concepts so that learners can work independently after explanations have been given. The independent work time is meant to be used as an opportunity for practice and consolidation of learning in groups or pairs.

Lessons learned:

- Training and support should emphasise the connection between practice and consolidation and the importance of adequate time being given to explanation and discussion of concepts prior to independent work.
- The TG does not include all of the steps that a teacher could follow to maximise the 45 minutes of the concept development session due to space constraints. Perhaps there is a need for a section that provides a few core strategies that teachers could choose from in order to optimise use of time.

While doing independent work, students are given opportunities to apply what they have learned, generally by doing activities in the LAB, facilitated by the teacher as their reading ability may not be at the required level to read the instructions independently. They also work with their peers during this time. Figure 7 shows that in LP and EC, students spent more time working independently (50% and 54% respectively) than in WC (26%).

Lessons learned:

- The videos show that teachers are allocating more time to independent work which aligns with beliefs we have heard teachers express in the field.
- Highlighting the importance and rationale behind time allocations is important both in training sessions and in support-visit sessions so that teachers do not compromise the learning opportunities of the learners.

Finally, Figure 7 shows that not much time is spent doing games, although games are an integral part of the BW programme.[8] We have observed in the field that some teachers find implementing games challenging. One of the main reasons they give for this is that there is not a written record for the time spent playing a game.

Lessons learned:

- There should be a weekly recording sheet for the games and a space for teacher's notes such as 'Learners who should play the game again on Friday' as a consolidation activity.
- More research needs to be conducted to understand the barriers to playing the games with the learners.
- Teachers could be encouraged to use the games as an early finishers' activity.

4.2.2 Integrity of implementation

The level of fidelity to the scripted lessons was coded under the variable 'integrity of implementation'. There were three categories: 'good' (following the TG), 'medium' (partial following of the script: some parts of the lesson are derived from the TG while other parts are not) and 'poor' (a lesson that had nothing to do with the TG guidelines). Findings for this variable are shown in Figure 10.

8 Note that only three of the observed lessons had scripted games which partly explains the low percentage.

Figure 10 Integrity of implementation

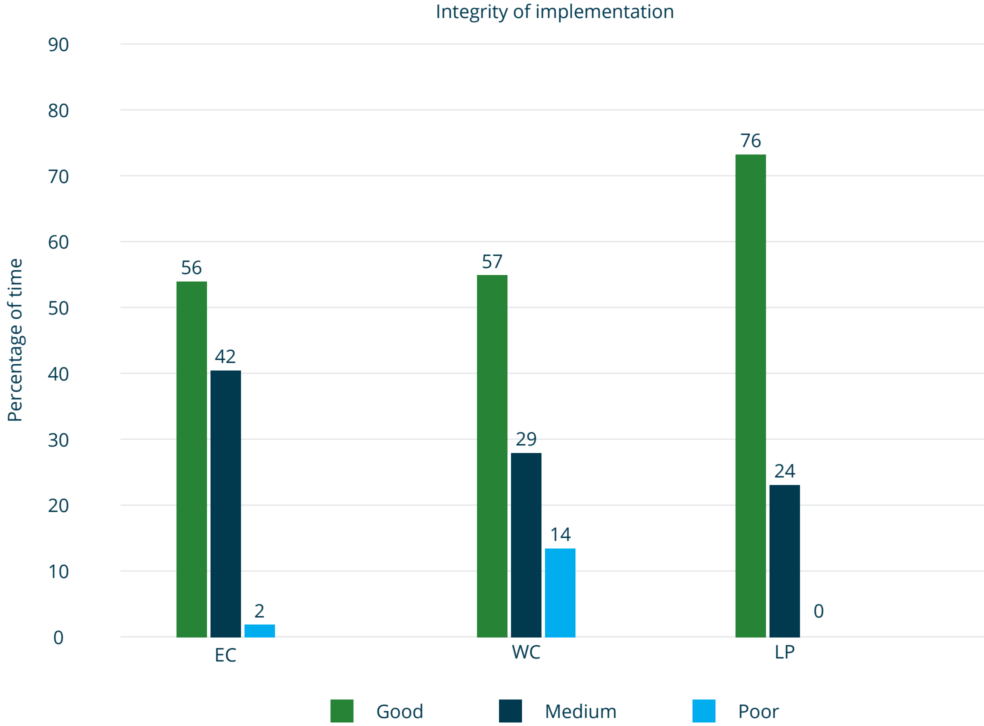


Figure 10 shows that there was a high level of integrity of implementation, more so in LP (76%) but also in EC and WC (56% and 57% respectively). This may have been an observation effect but even if it was, the figures show that teachers are able to interpret the scripted lessons. Only a few teachers from the sample were seen to disregard the TG, more so in WC (14% in the WC compared to 0% and 2% in other provinces). While compliance may be attributed to monitoring and support from the mathematics team, it would seem that at least some teachers are embracing the programme. One of the reasons for this difference may be attributed to the amount and type of training and support given in each of the provinces.

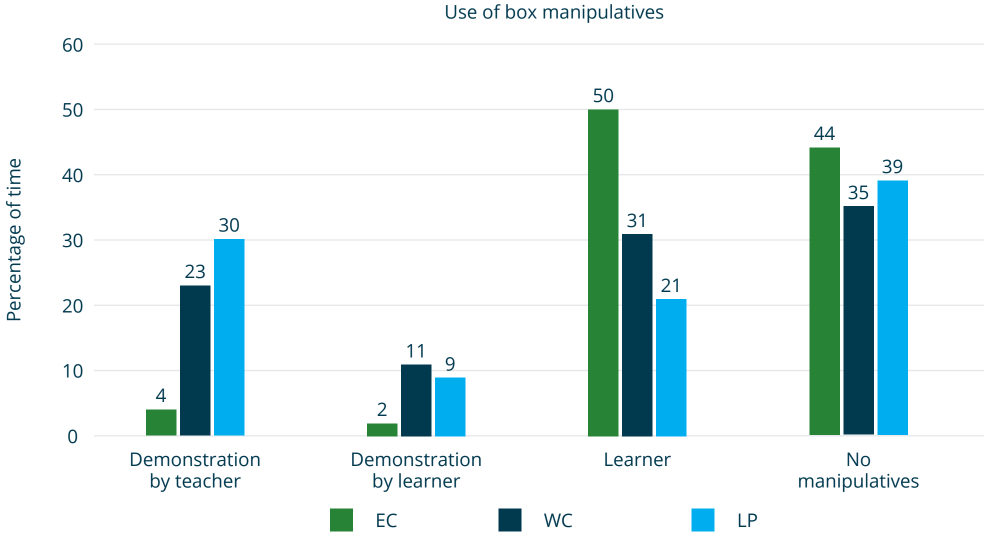
Lessons learned:

- Sufficient training is essential to support teachers' understanding of the programme and content. According to the time agreed to in each province, training times vary. LP training is significantly longer than in WC (two days compared to two hours per term). EC training is on an ongoing basis.

4.2.3 The use of manipulatives

There were three variables that captured use of manipulatives. We observed the use of manipulatives for counting, addition, subtraction and demonstration of the part-part-whole table to show number bonds. Figure 11 shows time spent working with the box manipulatives according to four categories.

Figure 11 Use of box manipulatives



It can be seen in Figure 11 that use of manipulatives varies across the provinces. Highest use of manipulatives by students ('learners') was in EC (50%), compared to LP (21%), and WC (31%). Demonstration using manipulatives by teachers was highest in LP (30%). Students did do demonstrations with manipulatives, but this was low across all provinces. There was a fairly high percentage of observed lesson time when no manipulatives were in use 35%, 39%, and 44% in the WC, LP, and EC respectively. It cannot be expected that manipulatives are in use 100% of the time during a lesson: it is more important that manipulatives are used; and that when they are used, they enable deep learning of concepts under discussion.

In the EC and WC use of manipulatives was also seen by students doing independent work.[9] This shows that teachers are encouraging students to work with manipulatives when they solve problems even though it is not scripted. The use of concrete objects in problem-solving can deepen understanding of concepts and build students' number sense especially for struggling students.

Lessons learned:

- The use of manipulatives during independent work should be emphasised during training to highlight the benefits observed as well as the different ways in which teacher agency has been used to strengthen the programme.

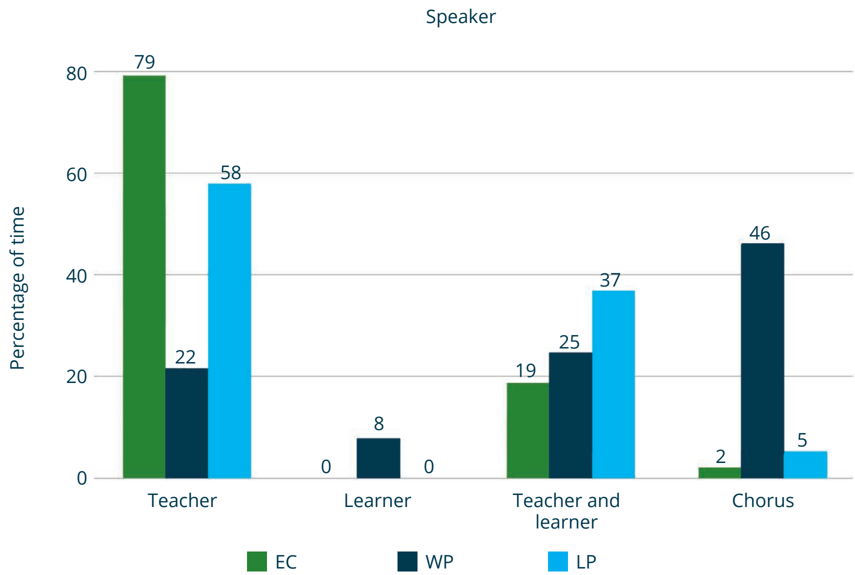
4.2.4 Talk in the classroom

Talk in the lessons was captured by the variable 'speaker' where four categories were noted. In a teacher-dominant segment, the teacher's voice is dominant, having a monologue in a form of an explanation, or modelling an activity. In a learner-dominant segment, students are speaking to each other about the activity they are engaged in, whether it be explaining the strategy they used, or working in pairs to solve a problem,

9 This was coded under another variable the graph of which is not shown here.

or peer teaching or explaining their strategy to the whole class. Teacher-and-learner-dominant would be characterised by the teacher asking a question, students responding to that question and the teacher acknowledging their answer (in this category, both voices are heard, but predominantly the teacher's). 'Chorus' is dominated by students reciting or recalling facts together. Findings for this variable are shown in Figure 12.

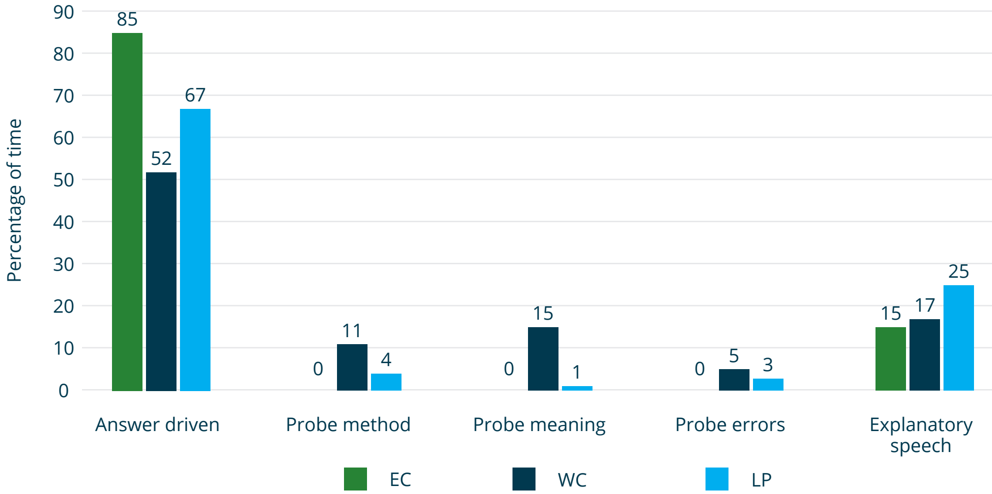
Figure 12 **Dominant speaker**



It can be seen in Figure 12 that dominant teacher-talk was most prevalent in EC (79%) but there was also evidence of 'teacher and learner' (EC, 19%) in the observed lessons. If we combine these categories for the EC, since they both represent teacher-dominant talk (Hattie, 2012), the percentage for dominant voice of the teacher rises to 98%. The remaining 2% is students giving chorus responses. There were no segments where student-dominant talk was observed. In LP 58% was dominated by the teacher and 37% by teacher and learner, a total of 95% of teacher-dominant voice, and 5% on chorus responses. Thus it can be seen that in these provinces the students' voices were not heard much. In the WC, the pattern was different and to a small extent, students' voices were heard. Teacher-dominant talk was 22%, teacher-and-learner-dominant talk was 25%, a total of 47% of teacher-dominant talk. Chorus accounts for 46% and 8% was student-dominant talk.

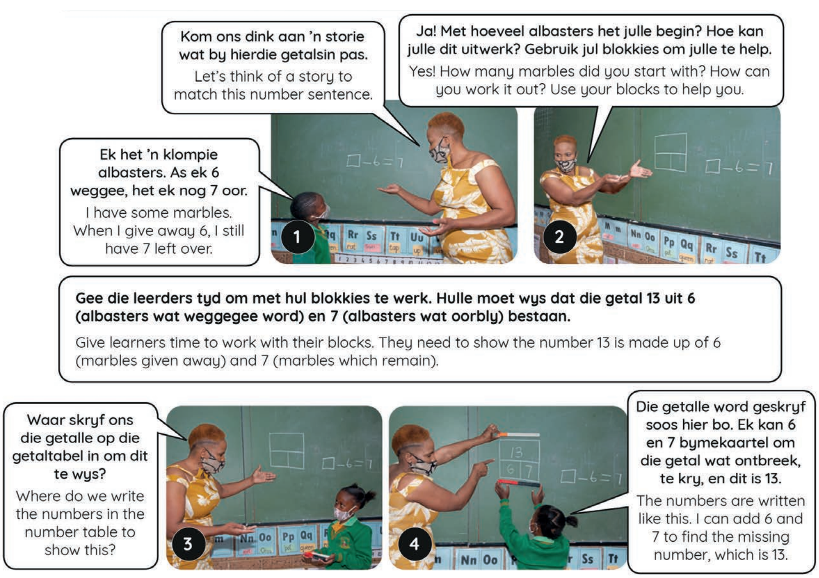
Our interest was engaging further with the types of dialogue that occurred during these lessons especially considering the concentration span of Grade 1 students. How did the teacher keep the students engaged? What types of questions stimulated the dialogue? The findings for the variable that addresses these questions are presented in Figure 13.

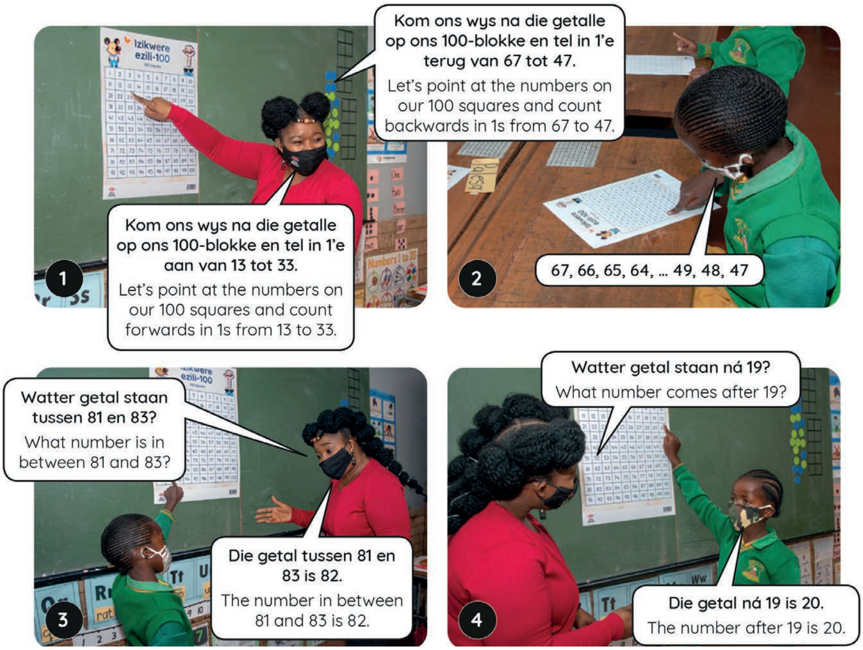
Figure 13 Questioning



It can be seen in Figure 13 that answer-driven questioning was most prevalent (52%, 67%, and 85% in WC, LP, and EC respectively). This category of questioning is characterised by low-level cognitive questions that require factual recall (Ellis, 1993). There were no high-level cognitive questions observed (probing method, probing meaning, or probing errors) in EC. The link between teacher-dominant talk and low-level cognitive questions is strongest in EC. Hattie (2012) argues that a classroom dominated by this type of engagement limits and discourages students talking together about their learning. An example of different kinds of questioning in the photo-scripts is shown in Figure 14 (both high-level and low-level questions are scripted).

Figure 14 Excerpts from the Bala Wande Teacher Guide (Grade 1 Term 4, pages 38 and 58)





Source: Bala Wande Grade 1 Afrikaans Teacher Guide Term 4, pages 38 and 58. From: https://fundawande.org/learning-resources?i=8#learningResourcesHolder

Lessons learned:

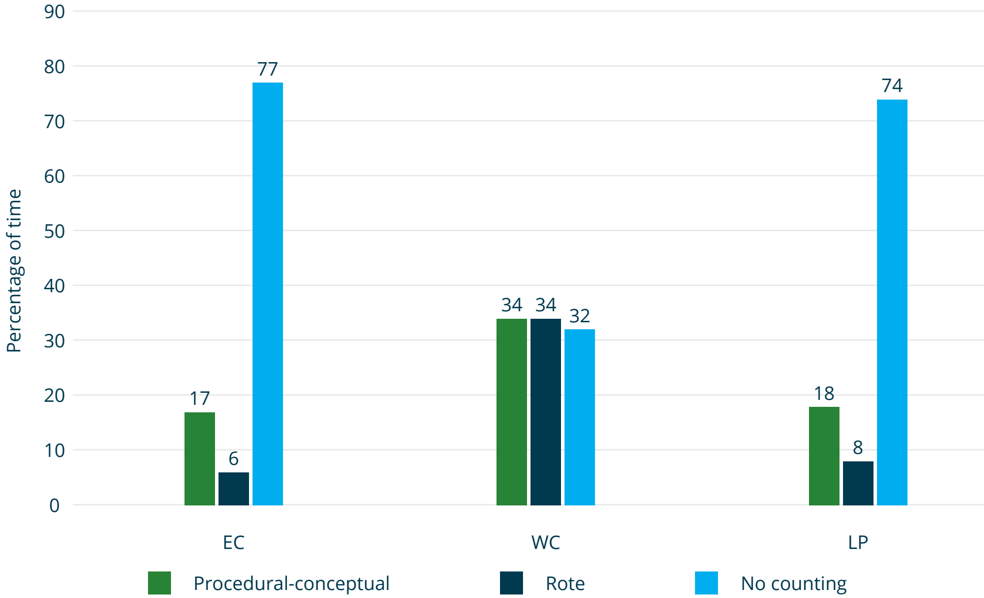
- The type of questioning modelled in the scripted photo sequences should be examined to check that it models sufficient high-level questioning.

4.2.5 Moving away from counting

Counting is considered an important activity in the development of number sense and the learning of mathematics in early grades. The BW goal in terms of counting is to reduce the amount of time that is spent on rote counting[10] activities in the mathematics classroom and to shift this to meaningful counting activities with a focus on the structure of number. Meaningful counting opportunities are created when a focus on number/quantity is provided in counting activities. Meaningful counting enables the development of number sense (it promotes the ability to instantly recognise the number of objects in a small group without the need to count them) and thus lays the foundation for working with additive operations using bonds rather than unit counting. In this study, there was a variable that captured three categories of counting activities observed during lessons. This is shown in Figure 15.

10 Rote counting is the ability to recite numbers from memory. Students recite numbers like a rhyme, with no attachment to meaning when they are rote counting. Rote counting is often the first counting activity of young students.

Figure 15 Counting



It can be seen in Figure 15 that there is considerable variation in the type of counting activities in the observed lessons across the three provinces.[11] The categories of procedural counting and rote counting are of interest here. In WC, the same amount of time (34%) is spent on procedural/conceptual counting and rote counting. The relatively higher percentage in WC for rote counting corresponds with the finding in Figure 4 that there was more time spent on mental maths activities in WC. In LP and EC, more time is spent on procedural/conceptual counting than rote counting (6% vs 17% and 8% vs. 18% respectively). The higher percentage of time spent on meaningful (procedural/conceptual) as opposed to rote counting in the provinces is taken as an indication that the programme goal of moving away from rote counting is beginning to be felt.

Manipulatives that were seen to be commonly used across the three provinces were: dot cards, 100-squares, ten-frames, beads strings, multifix blocks, number cards, number name cards, number symbols and number lines.[12] These manipulatives were relevant to the observed lessons and all of them (except the number name and symbol cards) are linked to the core goal of the BW programme of moving students away from counting by focusing on the structure of number.

Lessons learned:

- There is potential to shift learners away from rote counting but not all teachers have made the shift. We need to consider ways in which this change can be strengthened in order to allow the development of number sense. This is critical if the programme goal of enabling the shift away from unit counting when working with operations is to be met.

11 'No counting' occurred generally when it was not expected that students would be counting – for example when they are working independently or listening to teacher's explanations, and so on.

12 This was coded under a separate variable not represented graphically here.

5 Conclusion

Teachers using the BW materials benefit from having access to pre-selected examples for both teachers and students. The BW programme aims to support, draw on, and develop teacher agency by offering material that teachers are free to use on their own or supplemented by other examples, according to their professional discretion. The essence of learning mathematics is to reach abstraction. The design of the materials is focused on enabling this move, through the inclusion of the photo-scripted lesson activities that can be used to guide teacher preparation and print material in support of independent learning as discussed above. We expected take-up in the high-agency teacher sample to be relatively high, and across the three intervention sites, teachers demonstrated a relatively high degree of fidelity to the programme materials.

Teachers are more likely to embrace an intervention if they perceive it as high-quality and having potential to add value. The quantitative findings indicate that there is some take-up of the programme and that there are already some positive effects. They also showed that in schools where there was more support (LTSM+TA arm in LP) the effect was stronger. The qualitative findings showed that the programme materials (print and manipulatives) are being used. The most highly used manipulatives were ten-frames and multifix cubes, indicating that students are being exposed to structured number representations and there is potential for the lesson activities (and manipulatives) to support the shift away from rote counting when developing number sense, which can be built on to support the use of number bonds, rather than unit counting when working with additive operations. The quantitative evidence shows strength in programme schools relative to control schools which may be an effect of the programme activities and materials. Although it is small, it is a sign of positive change that could grow over time. Findings thus begin to address the question of what can be scaled, although more time is needed to gauge the full effect and consider all the implications for programme review. They also provided us with themes that need to be addressed during training sessions and inform materials development and review.

References

Amnesty International. (2020). South Africa: Broken and unequal: *The state of education in South Africa*. amnesty.org.

Aunola, K., Leskinen, E,. Lerkkanen, M. & Nurmi, J. (2004). Developmental Dynamics of Math Performance From Preschool to Grade 2, Journal of Educational Psychology, 96,(4), 699–713.

Brown, M., Askew, M., Baker, D., Denvir, H., Millet, A. (1998). Is the National Numeracy Strategy research based? *British Journal of Educational Studies*, 46(4), 362–385.

Brualdi, A.C. (1998). Classroom questions: ERIC/AE Digest, ERIC Digest Series No. EDO-TM98–02, Los Angeles, CA: ERIC Clearinghouse for Community Colleges.

Christiansen, I., Bertram, C., & Mukeredzi, T. (2018). Contexts and concepts: analysing learning tasks in a foundation phase teacher education programme in South Africa, *Asia-Pacific Journal of Teacher Education*.

Clements, D. H., & Sarama, J. (2014). Learning and teaching early math: The learning trajectories approach. Oxon: Routledge.

Department of Basic Education. (2018). Mathematics Teaching and Learning Framework for South Africa, Teaching Mathematics For Understanding. Pretoria: DBE.

Ellis, K. (1993). Teacher questioning behavior and student learning: What research says to teachers. Paper presented at the 1993 Convention of the Western States Communication Association, Albuquerque, New Mexico. ERIC Document Reproduction No. 359 572.

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), a481.

Gersten, R., Beckmann, S., Clarke, B., Foegen, A., Marsh, L., Star, J. R., & Witzel, B. (2009). *Assisting students struggling with mathematics: Response to Intervention (RtI) for elementary and middle schools* (NCEE 2009-4060). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Available at: http://ies.Ed.gov/ncee/wwc/publications/practiceguides/.

Griffin, C.P. (2017). Reflections on Using Video as a Data Collection Tool in Narrative Inquiry. Travel and Tourism Research Association: Advancing Tourism Research Globally. 1. Available at: https://scholarworks.umass.edu/ttra/2017/Qual_Methods_Papers/1.

Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning*. New York, NY: Routledge.

Heath, C., Hindmarsh, J., & Luff, P. (2010). *Video in qualitative research: Analysing social interaction in everyday life*. London: Sage Publications Ltd.

Howie, S.J., & Chamberlain, M. (2017). Reading performance in post-colonial contexts and the effect of instruction in a second language. IEA policy brief series. No 14. May 2017.

Jordan, N. C., Kaplan, D., Oláh, L.N., & Locuniak, M.N. (2006). Number Sense Growth in Kindergarten: A Longitudinal Investigation of Children at Risk for Mathematics Difficulties. *Child Development*, 77(1), 153–175.

Marra, M., & Holmes, J. (2008). Constructing ethnicity in New Zealand workplace stories. Text *and Talk – An Interdisciplinary Journal of Language, Discourse Communication Studies*, 28(3), 397–419.

Meehan, H. (1979). *Learning lessons*. Cambridge, MA: Harvard University Press.

Molina-Azorín, J.M., & Cameron, R. (2010). The Application of Mixed Methods in Organisational Research: A Literature Review. *The Electronic Journal of Business Research Methods*, 8(2), 95–105

Morgan, P.L., Farkas, G., & Wu, Q. (2009). Kindergarten predictors of recurring externalizing and internalizing psychopathology in the third and fifth grades. *Journal of Emotional and Behavioral Disorders*, 17, 67–79.

Pugalee, D.K (2005). *Writing to develop Mathematical Understanding*. Norwood, MA: Christopher-Gordon.

Resnick, L.B.C., Asterhan, C.S.C., & Clarke, S.N. (2015). Introduction: Talk, Learning, and Teaching. In: L.B. Resnick, C.S.C. Asterhan & S.N. Clarke (Eds) *Socializing Intelligence through academic talk and dialogue*, 1–12. Washington, DC: AERA.

Sowa, P., Jordan, R., Ralaingita, W., & Piper, B. (2021). Higher Grounds: Practical Guidelines for Forging Learning Pathways in Upper Primary Education. RTI Press.

Spaull, N. (2013). Poverty & Privilege: Primary School Inequality in South Africa. *International Journal of Educational Development* 33, 436–447.

Urquhart, C. (2013). The contribution of gtm – some reflections. In *Grounded theory for qualitative research*, 176-190. SAGE Publications, Ltd.

12

Bringing the Mental Starters Assessment project to scale in Foundation Phase: A 'building your timber' approach

HAMSA VENKAT & MELLONY GRAVEN

Abstract

In this chapter, we share details of a South African mental mathematics intervention (the Mental Starters Assessment Project) focused on Grade 3 learners. The intervention was developed over a six-year period in a model that went iteratively from small-scale local trials in two provinces to provincial-level trials into national-level implementation. This measured 'unhurried' model of building to national-level implementation is described in the chapter as a 'building your timber' approach, because of its attention to developing the capacity for scaling up in the course of scaling up. We comment on the ways in which the 'build your timber' approach contrasts with another scaling model seen in South Africa, described in the chapter as a 'straight for scale' model. In this latter model, interventions are designed for large-scale implementation from the outset. Overlaps and contrasts in the theories of change, the literature bases drawn from, the foci of teacher development, and the costs and timelines of implementation and research are considered in this chapter.

KEYWORDS
Mental mathematics, diagnostic assessment, fluencies, strategies, early number, scaling up, RCT, South Africa

1 Introduction

In this chapter, our focus is on the process of designing and scaling up an intervention – the Mental Starters Assessment Project (MSAP) – focused on improving the teaching and learning of mental mathematics in the Foundation Phase. The reasons for focusing on mental mathematics have been well argued in South African research. This research points to the persistent prevalence of counting-based approaches to number work well into the middle grades (Schollar 2008; Graven et al. 2014), coupled with evidence of early grades teaching that holds children back into counting quantities out in ones even when they show awareness of knowing the answers as recalled or derived facts (Ensor et al. 2009; Venkat & Naidoo 2012).

But problems in the South African education system extend beyond lags in learning and lack of attention to progression in teaching. Capacity issues have been identified in relation to instructional leadership in schools (Mestry, 2017) and in terms of district-office- and provincial-level resources to support the improvement of mathematics teaching (Metcalfe & Witten 2019). These multiple-level challenges have implications for thinking about interventions that can work at scale, and how interventions can be brought to scale.

In this chapter, we discuss key features of the MSAP scaling-up process in relation to an approach that we characterise in terms of 'building your timber'. Building your timber (BYT) refers to an approach wherein the need to build capacity for roll-out of an intervention is factored into the intervention design. A key feature of the BYT model is an acknowledgement of the need to build capacity within the system in the process of scaling for the focal intervention's roll-out. This entails starting small-scale and building up iteratively to scale in a design research model where evidence of positive learning outcomes is a prerequisite for scaling up from one stage to the next, and implementation experiences of stakeholders in the system inform the initiative's model in its next scaling-up iteration. In the chapter, we comment on the ways in which the BYT approach contrasts with an approach that we describe as the 'straight to scale' or STS model. Initiatives on the STS side tend to be large from the outset, working with 100 schools or more. A range of other differences tend to dovetail on either the BYT or the STS sides. In this chapter, we note some of the overlaps and contrasts between the MSAP as an example of an initiative using the BYT approach, and other initiatives that have used some elements of what we see as key markers of STS models. While initiatives often draw from elements of both of these models, highlighting their overlaps and differences allows for better understanding of the ways in which change in early grade mathematics (EGM) learning outcomes might be supported from a local research base, and identifying features that appear critical to consider in EGM initiatives going forward.

2 The MSAP

The MSAP began formally in 2016, with antecedent work in two South African Numeracy Chair research and development projects that have been working on early number-learning since 2011. The project's initial implementation was led by the two

Chair teams in local trials in a handful of schools in two provinces in 2017. Following positive learning-gains data across several iterations that grew in scale, the Mental Starters Assessment was included in national policy roll-out in 2022. The outcomes of the early trials have been detailed in Graven & Venkat 2021. Our attention here is on the model of scaling, the rationales for it, and our reflections on how this scaling-up model differs in important ways from 'straight to scale' interventions.

Below, we outline key aspects of the MSAP intervention content, focus, and model, weaving in the underlying literature that provided the rationale for our choices. We then detail and discuss the BYT approach highlighting key points of overlap and contrast with STS approaches.

2.1 The MSAP intervention content

The MSAP intervention was developed as a response to ongoing South African evidence of problems with using number-sense-oriented approaches to calculation in the primary grades. Schollar's (2008) graphic images of children in the later primary years working on two-, three- and four-digit multiplication and division problems using pages filled with counting-based approaches – either tallies counting out quantities in ones or repeated addition-based counting in multiples – illuminate one facet of the problem. Our research with partner schools in Gauteng and the Eastern Cape showed similar tally-based strategies even for simple addition and subtraction problems such as 10 + 10 + 10 and 23 + 18 (see Weitz & Venkat 2013 and Graven & Stott 2012 respectively).

In a context where schools usually provide the sole sites of access to formal mathematics learning, learners' lack of acquisition of number sense reflects limited focus in mathematics teaching on mental mathematics underpinned by a strong awareness of number relationships and how the structure of the base-ten decimal system can leverage efficient calculation. Interventions in both of our projects had attended to different packages of materials and teacher development to address these issues – focusing on learning lags via working with teachers to improve their awareness of strategies leveraging these efficiencies, and on supporting their teaching for inclusion of attention to these efficiencies. In the work of Wits Maths Connect-Primary (WMC-P) project involving the first author, the Structuring Number Starters project had focused on efficient calculation and 'base-ten thinking' in a school-based package of termly workshops for Foundation Phase teachers and teacher and learner mental-starter materials for use in the Curriculum and Assessment Policy Statement (CAPS)-stipulated mental-starter section of lessons. CAPS suggests all lessons start with whole class brisk mental starters such as "the number before 8 is; 2 more/less than 8 is; 4+2; 5+2; 6+2 etc". (DBE 2011a, 9). The policy further suggests a range of mental strategies that learners should develop including the six we focus on in the MSAP, namely: bridging through ten, jump strategy, doubling and halving, understanding the relationship between addition and subtraction, re-ordering, and compensation (DBE, 2011a).

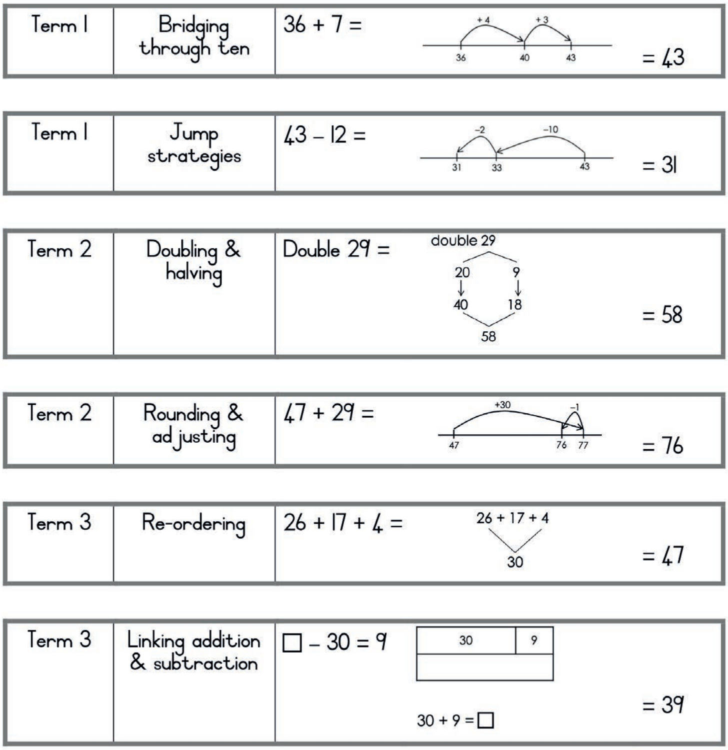
In the Rhodes University-based South African Numeracy Chair (SANC) work led by the second author, the poorer and less functional schooling system in the Eastern Cape fed into the development of models, materials, and assessment packages focused on mental maths fluencies for use in classrooms and after-school Maths Clubs (see Graven & Stott 2012; Stott & Graven 2013). Both projects, over time, had amassed evidence of positive impact on learning outcomes – in schools in the WMC-P project and for use in classrooms and after-school Maths Clubs in the SANC project. In the WMC-P project, there was evidence of increases over time in the proportions of early Grade 2 learners able to work with efficient mental and/or written methods that made use of base-ten structure (Venkat, Askew & Morrison, 2021). These methods involved, for example, calculating the answer to addition and subtraction problems involving two-digit numbers mentally either by splitting both numbers into tens and units and then recombining for the final result, or taking the first number as a starting point and 'jumping' forwards or backwards by the required number of tens and units. In the SANC project, there was evidence of improvements in fluency measures and using more efficient and less error-prone strategies by Grade 3 and 4 learners in partner schools and after-school clubs, as indicated by timed fluency and four operations assessments (see Graven 2015 for data across partner schools, and Stott et al. 2017 for data across thirty clubs in three provinces).

We disseminated these early results from partner schools in a range of national fora, including DBE seminars, the annual Mathematics Chair Community of Practice forums and the annual conferences of the Association of Mathematics Educators of South Africa (AMESA) and the Southern African Research Association for Mathematics, Science and Technology Education (SAARMSTE). These are key platforms where teachers, researchers, teacher educators and DBE curriculum and assessment personnel come together to grapple with ways forward to challenges in mathematics education in South and southern Africa. Dissemination of our findings in these platforms and rich dialogue with multiple stakeholders led to discussions with the Department of Basic Education (DBE) in 2015 on developing materials and assessments for Foundation Phase mental mathematics. The DBE Assessment Unit had noted the dissatisfactions expressed by teachers with the summative Annual National Assessments, including the teaching time given over to preparing learners for assessments (see Graven & Venkat 2014), and were looking to meet teachers' requests for low-stakes, in-class assessments, that would offer more of a 'connect' rather than 'disconnect' with teaching. In bringing together this request with our prior experiences on working on mental mathematics underpinned by number sense, the MSAP was born.

2.2 The MSAP's mathematical focus

In 2016, the Wits and Rhodes Chair teams came together with two international early-number experts (Professors Mike Askew and Bob Wright), the DBE, provincial and district personnel, and representatives from the Association of Mathematics Educators of South Africa (AMESA) and SAARMSTE. The work in the initial meeting was to consider priorities for mental mathematics working that would align with the CAPS curriculum and the international literature base on early number knowledge. The meeting led to two key decisions – the first relating to what mental mathematics strategies to focus on, and the second relating to how to focus on these strategies. In terms of the 'what', six strategies were selected for use across Grade 3, which came to be labelled and illustrated in the teacher materials that were completed in 2020 (Graven et al. 2020a and b) and then versioned into all eleven official South African languages by the DBE (DBE 2021a and b) as shown in Figure 1:

Figure 1 **The six mental mathematics strategies in focus in Grade 3 (DBE 2021a, page iii)**



Source: Mental Starters Assessment Project English Teacher Guide, page iii. From: https://www.education.gov.za/MSAP2022.aspx

Our decision on how to focus on these strategies in instruction was based on Askew's (2012) emphasis on fluencies, reasoning and problem-solving as the 'visible' aspects of Kilpatrick, Swafford & Findell's (2001) earlier writing on the strands of mathematical proficiency. Attention to fluencies was intended to address the need for fluent recall of a basic bank of number facts: bonds of single-digit numbers, doubles of 1-10, adding or subtracting multiples of ten from a number, and knowing the size of jumps from given numbers to the next or previous multiple of ten among these. The literature base on early number describes efficient problem-solving as resting on a growing bank of fluencies – for example, calculating 46 + 25 mentally in efficient ways might involve working this sum out as 46 + 20 = 66, then 66 + 4 = 70, and finally 70 + 1 = 71. Adding multiples of ten to any number, awareness of the next multiple of ten, and how to jump from 66 to 70 in this case, and knowing the associated bond of 5 (i.e. 4+1) are all fluencies underlying this kind of 'strategic' problem-solving. These approaches are more effective and more efficient than the laboured effort and error-prone approaches of counting in ones for this problem.

Reasoning refers to awareness of the number relationships underlying such calculations, rather than to the calculations themselves. It is the structure of additive relationships and the properties of the decimal number system that are in focus in this thread, seen in tasks such as: 99 + 99 = 200 – __ , or in the case where a given result can be used to generate a range of other expressions, including expressions based on inverse operations. In the mathematics education literature, this kind of reasoning is seen as critical to flexible working with missing number problems. An example of this is seeing:

$$31 - \_\_ = 16 \text{ and } 31 - 16 = \_\_ \text{ and } 16 + \_\_ = 31$$

as interlinked problems rather than as separate and unconnected problems (Baroody 1987). Further, being able to reason about the structure of number relationships in missing number problems is an important foundation for later work with algebra, an area where South African evidence points to significant problems for older learners (Pournara 2020).

2.3 The MSAP intervention model

The MSAP intervention model involved a sequence of three activity segments for each of the mental strategy-focused units highlighted above. These were:

- A 5-minute written test, set on a Monday morning by the Grade 3 class teacher, in the mental-starter section of mathematics stipulated for lessons in the CAPS curriculum. The test includes fluency items on one side to be completed in two minutes, and reasoning and problem-solving items on the reverse side to be completed in three minutes.
- Teaching of eight mental-starter activities within a three-week block following the pre-test. Each starter includes attention to fluencies in a short warm-up section, and is followed by a combination of teacher-led work and individual work on reasoning and problem-solving tasks related to the focal strategy. Two practice worksheets for learners to take home are included for use across this period.
- A five-minute written post-test, set on the Friday morning at the end of the three-week period by the class teacher, again in the mental-starter section.

The tests and outline of teaching activities are detailed in the MSAP Teacher Guide document (DBE 2021a). A separate MSAP Print Masters document collates all the ‘printables’ (tests, worksheets, and flash cards) in a separate book for easy copying (DBE 2021b). Marking pre- and post-tests linked with each unit provides the teacher with feedback on areas of efficacy in her teaching, and a sense of where limitations remain in learner working.

As detailed in the outline above, two mental mathematics units are scheduled for each term for use in the starter section of lessons. The aim is to build early number competence, through careful attention to key representations that permeate several primary mathematics topics (e.g. number lines, part-whole diagrams, doubles images) and to task variations that can feed into learners’ work in the main lesson section across a range of topics in the mandated curriculum that includes number sense. The assessment model is low-stakes, and oriented towards a focus on supporting responsive and diagnostic classroom teaching rather than a focus on systemic monitoring. In this respect learners and teachers are encouraged to shift their focus from raw scores on tests to improvement of scores from the pre- to post-test. Our experience is that this focus on improvement rather than scores reduces some of the anxiety associated with testing that is time-sensitive.

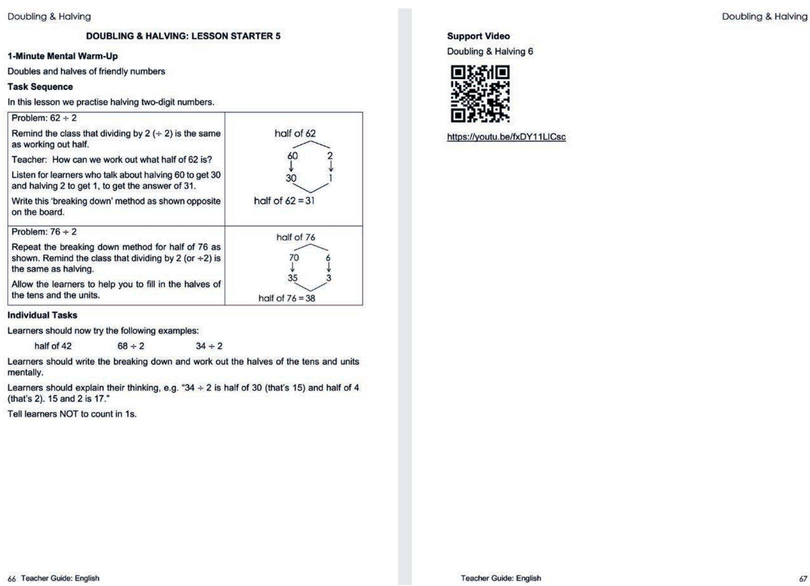
3 Taking MSAP to scale: The BYT approach

Following the initial design and development of assessment and teaching materials for the early units in 2016, both teams undertook initial trials in our 'home' provinces. In Gauteng, WMC-P team members trialled the Bridging through Ten unit in two schools: one a quintile 1 township school and the other a quintile 5 suburban school. In the Eastern Cape, the SANC team trialled the same unit in a quintile 3 school. In both provinces, these trials involved the research team working directly with the classroom teacher, explaining, discussing and demonstrating both the materials themselves and the mathematical rationales for using the tasks, resources and representations in particular ways.

These trial results are written up elsewhere (Graven & Venkat, 2021), so we do not discuss the outcomes here. Instead, we note simply that the outcomes pointed to a proof of possibility – that the materials could be used by teachers in the system to produce learning gains. But scaling up the intervention necessarily involved going beyond our research teams as the trainers. Following Borko's (2004) iterative model for scaling-up interventions beyond the research teams where they originated, we turned to the Foundation Phase subject advisors, with both projects having done prior developmental work with subject advisors in their own provinces. In the WMC-P project, key shifts in the orientations of subject advisors towards mathematics and towards collaborative, rather than monitoring/evaluative relationships with teachers, had been noted in their earlier work (Venkat & Askew 2021). In the Eastern Cape DBE teacher educators (e.g. Foundation and Intermediate Phase subject advisors, Foundation Phase curriculum planners and district cluster leaders) were a key part of the professional development programmes and key to rolling out the clubs in a range of districts across the province. This prior groundwork pointed to possibilities for working with personnel at the district level in the South African education system, contrasting the evidence noted earlier of limitations in the capacity of this layer to contribute to development.

Collaborating with the DBE in 2019 allowed for a three-province trial that took in the range of low-, mid- and high-performing provinces in the country. The research question driving this first level of scaling-up was to understand the extent to which the intervention, now mediated by district subject advisors (supported in turn by the WMC-P and SANC research teams), could produce pre- to post-test gains on two units. The results, also reported in the aforementioned paper, suggested that gain levels could be maintained in working via the district level infrastructure in the system. This result was important, as it suggested that – with careful attention to supporting materials that personnel in the system could see the need for – the level of pre- to post-test improvement seen in small-scale trials could be maintained. Materials were shared and discussed in two carefully-developed training sessions with interim monitoring and support for roll-out from the DBE, A trial carried out in 2021 involving all the Foundation Phase mathematics subject advisors in the country (see Askew et al. 2022 for data from the trial) led to further important developments. Firstly, broadening to all nine provinces involved the creation of materials for all six of the mental maths strategies, ensuring consistency of format of the assessments and of the mental-starter structure of working. In considering the additional supports needed when working with intermediaries to translate an intervention into practice, short 'talking hand' video clips of the central focus of each starter-activity were narrated and enacted by Mike Askew. These video clips were made accessible via both smartphone accessible QR codes and web addresses created by Lynn Bowie, an associate of the Wits team with experience of creating cellphone-accessible materials in her work in the NGO sector (see Bowie et al. 2022). This approach was chosen based on the evidence of widespread access to smartphones in the South African population (several surveys point to around 95% population access). It was also responsive to the finding of limited use of lengthy teacher guides and support materials among South African teachers. An example of a QR code linked to one of the starter tasks in the Teacher Guide booklet (DBE, 2021a, 66-67) is reproduced in Figure 2.

Figure 2 **Task with QR-code accessible video: Doubling and Halving, Lesson Starter 5**


Doubling & Halving

DOUBLING & HALVING: LESSON STARTER 5

1-Minute Mental Warm-Up

Doubles and halves of friendly numbers

Task Sequence

In this lesson we practise halving two-digit numbers.

Problem: 62 ÷ 2 Remind the class that dividing by 2 (÷ 2) is the same as working out half. Teacher: How can we work out what half of 62 is? Listen for learners who talk about halving 60 to get 30 and halving 2 to get 1, to get the answer of 31. Write this 'breaking down' method as shown opposite on the board.	half of 62 60 → 30, 2 → 1 half of 62 = 31
Problem: 76 ÷ 2 Repeat the breaking down method for half of 76 as shown. Remind the class that dividing by 2 (or ÷2) is the same as halving. Allow the learners to help you to fill in the halves of the tens and the units.	half of 76 70 → 35, 6 → 3 half of 76 = 38

Individual Tasks

Learners should now try the following examples:

half of 42   68 ÷ 2   34 ÷ 2

Learners should write the breaking down and work out the halves of the tens and units mentally.

Learners should explain their thinking, e.g. "34 ÷ 2 is half of 30 (that's 15) and half of 4 (that's 2). 15 and 2 is 17."

Tell learners NOT to count in 1s.

66 Teacher Guide: English

Doubling & Halving

Support Video

Doubling & Halving 6

https://youtu.be/fxDY11LICsc

Teacher Guide: English 67

Source: Mental Starters Assessment Project English Teacher Guide, pages 66-67. From: https://www.education.gov.za/MSAP2022.aspx

On the DBE side, this expansion further involved versioning the materials into all eleven official South African languages, working with the two Chair teams.

In this trial, each subject advisor worked with two or three Grade 3 teachers from one school in their districts on one of the first four units in the sequence. Different provinces worked on different units, with the two Numeracy Chair teams running a half-day training session at the start of the trial. The training sessions, as in the earlier trials, were focused on the mathematical content of the unit, with an emphasis on key tasks, representations, and explanations around using these in the starter section of lessons. This emphasis on teaching the mathematical content and skills related to the unit, with illustration in the video, was based on the evidence of teachers' difficulties with coherent and connected explanations in South African primary classrooms (Askew et al. 2019) and broader evidence of limited attention to progression in instruction outlined earlier. Covid-19 meant this training occurred online, hosted by the DBE on the Microsoft Teams platform. The advantage of this is that the recordings of the training are available for reviewing and for future use in training.

The DBE's involvement in the 2021 trial also meant that updates on the progress of the MSAP roll-out by subject advisors were retained as a standing item in the DBE Assessment Unit's regular meetings with provincial Foundation Phase coordinators. Our participation in these meetings has allowed for updates and reflections on progress with implementation. For example, we heard from the provincial coordinators about challenges of Covid-related school closures in some provinces soon after the start of the trial, but also about measures taken by some subject advisors to restart and complete the trial in other schools that remained open. We also gained insights into the number of districts that had managed to complete the trial, and feedback received from subject advisors who had either looked through pre- and post-test results, and in some cases, also observed and supported the administration of these tests in the classes of the teachers they were supporting. Responses from one province at these meetings noted, for example, the following feedback:

> *Learners enjoyed the Fizz Pop activity and would play it during breaks.*
>
> *During post-tests, learners' performance and speed improved tremendously in majority of learners.*
>
> *Use of number line enabled learners to bridge through ten in a simple and logic way.*
>
> *The use of number line enabled learners to identify the next ten from the given number.*
>
> *Learners managed to see the pattern of tens.*
>
> *Learners managed to develop their variations of the strategy.*
>
> *Bigger numbers challenged most of the learners.*
>
> *Learners took a few days to master the skill of subtraction in the number line.*

The specificity of this feedback was of interest to us, as it suggested a level of appropriation at district subject advisor level, of the mathematical goals of the intervention, and the ways in which key representations included in the materials could support moves towards these goals.

The 2021 trial was viewed as serving two key functions – firstly, it allowed us to understand the feasibility of roll-out in all provinces, and secondly, it served as a 'familiarising' avenue for the provinces coming on board for the first time. Results from the 2021 trial showed medium to large effect sizes across all the trialled units (see Askew et al. 2022 for details). In the midst of Covid-related disruptions to schooling, subject advisors from seven of the nine provinces were able to collate and submit matched pre- and post-test data from n = 2,294 learners that encompassed three of the four of the MSAP units trialled. The two remaining provinces also had Subject Advisers who had worked with teachers on their focal unit, but Covid-related school closures disrupted post-test administration, whilst maintaining the goal of familiarisation.

Within the BYT approach, an important aspect of the MSAP model is its three-week unit frame and its narrow focus on one particular mental mathematics strategy in each unit. These two features combined to offer the subject advisors a chance to see the impacts of their own training with teachers on one unit at a time in terms of learning gains, in ways that are harder to do when the focus is on the whole curriculum. In an education context where initiatives are frequent, where Covid has ravaged access and exacerbated inequality (Soudien, Reddy & Harvey, 2021), and where large-scale interventions offer infrequent insight beyond overall score-level changes in learner performance (in the instances of statistically significant gains in comparison to control groups – which is not always the case, and discussed below), qualitative illustrations of what children's improvement looks like are important. This is particularly so given the broader public discourse that repeatedly reinforces deficit narratives of poor performance in mathematics in national tests and in international comparative test outcomes (Shay 2020), which can feed into low expectations and a sense of inevitability of poor performance among teachers in poorer schools (Graven 2014). Our point here is not to dispute the realities of low performance in the majority of South African schools. Rather, we note that there is a sense of hopelessness in the iteration of this narrative that renders its mark on the faith that personnel within the education system have on the potential of interventions, and of their own work with interventions, to effect change. Evidence of learning gains, with a sense of how children are able to improve and develop mathematical fluency and reasoning over time, provides an important counter to this narrative.

4 Distinguishing features of BYT approaches

We began this chapter by pointing out that measured, unhurried, and deliberate design was a key feature of the BYT approach to building capacity within the system for taking interventions to scale in the course of roll-out. This contrasts with STS approaches where speedier roll-out involves testing the system's capacity for taking on the intervention. The larger-scale projects such as Programme to Improve Learning Outcomes (PILO), Gauteng Primary Language and Mathematics Strategy (GPLMS) and Bala Wande, align in different ways across the BYT and STS categories on this issue: the MSAP study's work with subject advisors and teachers places it firmly in the BYT camp; PILO similarly was predicated on the need to build capacity within the system, and also worked their interventions via subject advisors. In contrast, GPLMS worked with external 'coaches', hired by the project to support teachers and teaching. Bala Wande began with an approach that compared support for the implementation of their materials in classrooms coming from externally-hired coaches (Eastern Cape), subject advisors (Western Cape) and school-leaver teaching assistants (Limpopo) – and also included intervention and control schools. However, the Coaches model has been dropped from their latest plans. Thus, there are moves towards the BYT approach in the

Bala Wande model.

In this concluding discussion section, we focus on key areas of overlap and contrast of STS and BYT models, drawing from the additional features of MSAP discussed above and using these to briefly consider the features of other key South African initiatives that have worked at larger scales. In doing this, we draw on international writing on the affordances and constraints of both STS and BYT scaling models. These areas encompass, in inter-related ways:

- Theories of change
- Emphases within the materials package, and the literature bases they draw from
- The focus of teacher development
- The evidence-base used to gather data on intervention efficacy
- Timelines and costs.

It is of interest that several of the larger-scale intervention projects in South Africa have tended to be led by education policy-oriented academics (e.g. Brahm Fleisch: GPLMS; Mary Metcalfe: PILO; Nic Spaull: Funda Wande). This contrasts with the MSAP and other BYT interventions, nationally and internationally, which have tended to be located within mathematics education departments and research teams. Where an intervention originates has consequences for the theories of change that inform the intervention. Fleisch et al. (2016), writing on the GPLMS, underpinned by the triple-cocktail combination of mechanisms, provide some useful pointers. Their description of what the scripted lesson plans within the package sought to achieve is stated thus:

> *The scripted daily lesson plans were intended to provide systematic, paced, and accessible lessons for the teachers to follow through the year. The underlying purpose of the lesson plans was to introduce and gradually institutionalise a repertoire of practices that would improve teachers' time on task and establish new daily and weekly routines*
>
> (2016, 159).

What is notable for us within this description is the absence of mention of any specifics related to mathematics (or language) and its teaching. This 'generic' focus on teaching and learning is reflected in the literature base that Fleisch et al. (2016) draw upon primarily to inform their work, located in the educational change field. The emphasis here is on changing the norms and routines of teaching in terms of expectations of coverage and pacing – reflecting some aspects of what McKay and Spaull (this volume) describe as a focus on the 'grammar of schooling'. De Clercq and Shalem (2019), writing about the theory of change embedded in the PILO intervention, noted a similar emphasis on coverage and pacing communicated through this project's provision of curriculum planners and coverage 'trackers', but acknowledged too that more transparent and collegial conversations about curriculum coverage between HODs and teachers had been fostered during implementation. In contrast, our work in MSAP and that of others working on intervention studies located in mathematics education have tended, predictably, to draw centrally from the mathematics education literature to devise the content of interventions. Changing mathematics instruction in quite specific ways linked to the mathematical content in focus is therefore core to the theory of change in the MSAP rather than a more generic focus on coverage.

The packaging of materials into some form of scripted sequence features in many South African EGM interventions – GPLMS, MSAP and Funda Wande among these. It is worth noting though that the rationales for scripting tend to vary somewhat, with the change literature cited by Fleisch et al. rationalising lesson plans as:

> *designed to reduce teachers' planning and administrative workloads and allowed them to concentrate on actual teaching, thus shifting the focus from interpreting the national curriculum to delivery of the curriculum*
>
> (2016, 159).

In our work, the need for scripting rested more – as noted earlier – on the evidence of gaps in primary teachers' mathematical content knowledge and difficulties in orchestrating coherent and connected instructional explanations. It also rested on the absence of widespread use of key conceptual representations such as the empty number line or part-part-whole diagrams to support sense-making. This emphasis on conceptual mathematical representations, mathematical fluencies and strategies, and coherent, connected mathematical teaching, is carried through into the professional development offered within the MSAP. In our work with subject advisors, rationales for a focus on mental mathematics were made explicit in the training. Subject advisors acknowledged the need to deal with the persistent use of counting which was familiar from their work on the ground. Discussion of the tasks, representations, and need for a move from teacher presentation and interaction into individual working with the strategy in focus in the unit were all emphasised and agreed upon. The timelines in the MSAP intervention are looser than in the GPLMS and Funda Wande programmes: subject advisors are advised that two units can be covered each term across Terms 1-3, with three weeks suggested for coverage of each unit. Our interest, made explicit to subject advisors for their work with teachers, is whether teachers' work with starters within the three weeks can produce gains for individuals and classes of learners from pre- to post-tests for each unit. We also encouraged subject advisors (and through them, teachers) to attend to the differences in learners' ways of working with problems as the underpinning elements of any gains they see. This kind of noticing of, and attention to, children's mathematical learning over time (as operationalised in individual learner improvement scores from pre- to post-test and associated shifts to increasingly efficient methods of calculation) is core to the orientation towards professional development in MSAP. Developing such noticing has widely been described as a key component of improving teachers' pedagogic content knowledge (Jacobs et al. 2010). We see this as especially important in a context where the mathematical capabilities of personnel at district levels is also questioned. The BYT approach acknowledges the shortcomings of teachers "mathematical knowledge for teaching" (Ball & Bass 2003) and looks to address them, as part of the process of building mathematical support capacity in the system.

A range of research methodologies have been used across the studies discussed in this chapter. In the MSAP iterations, as already noted, gains were studied at the unit level with further breakdown within the units to look at fluency, strategic calculation and strategic thinking level gains. We have used pre- and post-test data, across the iterations of the intervention over time to study the learning outcomes of teachers working with the MSAP model and materials. As noted, GPLMS and Bala Wande included the use of intervention and control groups to compare the impact of their materials or training model on learning. We have seen less reporting of learning outcomes – whether in experimental or longitudinal designs – available for the PILO and NECT studies than for the aforementioned studies. This makes it quite difficult to gauge impact on mathematical learning outcomes, and represents a limitation of these two studies for understanding the possibilities and limitations of STS designs.

In the MSAP model, the absence of parallel control schools leaves open the critique of lack of counterfactual evidence – the fact that there is no data telling us about the learning gains in schools that did not participate in the MSAP trials, which may, or may not, have been similar to the kinds of gains seen in intervention schools. However, what our dataset does include are the qualitative-feedback elements exemplified earlier that offer specific details of what children are doing differently in their work with core mathematical tasks and representations, alongside specific aspects of what children are still having difficulty with. This noticing of distinctions with mathematical specificity has been noted as weak in South African research with subject advisors (Venkat & Askew 2021) and is therefore an important part of capacity-building.

Our final point relates to timelines and costs. The large scale of STS models and the inclusion of control schools inevitably makes this research expensive, and especially so given that the scale often requires the recruitment of researchers for gathering data. In intervention models involving additional personnel such as coaches, the costs further increase, threatening the sustainability of the intervention beyond the lifetime of the research. But critiques of the randomised-control trial component within STS approaches point also to the relatively common occurrence of the absence of treatment effects in these studies (e.g. Hedges & Schauer 2018). Advocates point out, and rightly so in our view, that null hypotheses findings are useful for understanding what does not work and for probing why. But these findings tend to be achieved through expensive mis-steps that the education system in South Africa can ill afford. This point makes the absence of learning outcome data in publications emanating from the NECT and PILO studies problematic – as we are left with no course for interpreting impact.

The MSAP model has been cheaply brought to national scale – supported by the five-year investment of our funders in our two numeracy research and development Chairs from which it emerged. Its cost-efficiency has resided in a model where small teams were involved in initial trials, and where scale-up worked with personnel in the system. They mediated and supported the trials with teachers as an extension of their everyday work. Negotiating their participation involved work, over time, with national and provincial education departments, alongside evidence from earlier trials and a commitment to the work of capacity-building. The DBE is currently exploring extensions of focus on mental mathematics related to the MSAP model across the primary grades 1-6, bringing further potential expansions of scale into view. Given the larger numbers of learners in STS approaches, the cost per learner might be relatively small, but we retain the point that STS studies remain significantly more expensive overall than BYT approaches, rendering the need for proof of the promise of the intervention to be much stronger at the outset. A caveat in BYT approaches is the timeline: direct work on the various MSAP iterations has taken six years to bring to the national policy level. STS models can often provide measures more quickly of the likely efficacy of an intervention, so we acknowledge the trade-offs.

In concluding, perhaps most fundamentally there is the sense that there is limited point in testing a fragile system for its potential to manage interventions and especially those that the system is unlikely to be able to afford beyond the research. The need to build substantive capacity within the system to support mathematically focused interventions is, for us, the urgent goal to focus on.

Acknowledgements

Our thanks to the DBE and the NRF for their support of this work. Thanks to our entire team who grappled together in developing these formats, assessments and reasoning chains namely Lawan Abdulhamid, Mike Askew, Nolunthu Baart, Lynn Bowie, Mark Chetty, Corin Mathews, Zanele Mofu, Samantha Morrison, Debbie Stott, Herman Tshesane, Lyn Webb, Marie Weitz, Bob Wright. Also, thanks to Lawan Abdulhamid, Sameera Hansa, Lynn Bowie and Pam Vale who worked with us on the data collection, collation and analysis.

Funding information

This work is based on the research supported by the South African Research Chairs Initiative of the Department of Science and Technology and National Research Foundation (Grant Nos. 74703 and 74658). Any opinion, finding and conclusion, or recommendation expressed in this material is that of the authors and the NRF does not accept any liability in this regard.

References

Askew, M. (2012). *Transforming primary mathematics*. Abingdon: Routledge.

Askew, M., Graven, M. & Venkat, H. (2022). From what works to scaling up: Improving mental strategies in South African Grade 3 classes. In Fernández, C., Llinares, S., Gutiérrez, A. and Planas, N. (Eds.) *Proceedings of the 45th Conference of the International Group for the Psychology of Mathematics Education*, July 18–23, Alicante, Spain. 27–34.

Askew, M., Venkat, H, Abdulhamid, L., Mathews, C., Morrison, S., Ramdhany, V. & Tshesane, H. (2019). Teaching for structure and generality: assessing changes in teachers mediating primary mathematics. In M. Graven, H. Venkat, A. Essien & P. Vale (Eds.). *Proceedings of the 43rd Conference of the International Group for the Psychology of Mathematics Education*, 41–48. Pretoria, South Africa: PME.

Ball, D. L. & Bass, H. (2003). Toward a practice-based theory of mathematical knowledge for teaching. In *Proceedings of the 2002 Annual Meeting of the Canadian Mathematics Education Study Group*. 3–14. Edmonton Alberta.

Baroody, A. J. (1987). *Children's mathematical thinking: A developmental framework for preschool, primary, and special education teachers.* Teachers College Press.

Borko, H. (2004). Professional Development and Teacher Learning: Mapping the Terrain. *Educational Researcher,* 33(8), 3–15.

Bowie, L., Collins, H., Mashiyane, N. & Matanhike, F. (2022) Maths clubs: growing the possibilities of the after-school space. In Venkat, H. & Roberts, N. (Eds), *Early Grade Mathematics in South Africa*. Cape Town: Oxford University Press.

De Clercq, F & Shalem, Y. (2019). Large-scale improvement interventions in the education system: PILO's contribution to the theory of change in education. *Southern African Review of Education*, 25(2): 34–51.

DBE (2021a). Grade 3 Mental Starters and Assessments: Teacher Guide (Creative Commons license) – all South African languages

DBE (2021b). Grade 3 Mental Starters and Assessments: Print Masters (Creative Commons license) – all South African languages

Ensor, P., Hoadley, U., Jacklin, H., Kühne, C., Schmitt, E., Lombard, A., et al. (2009). Specialising pedagogic text and time in foundation phase numeracy classrooms. *Journal of Education*, 47, 5–30.

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

Graven, M. H. (2014). Poverty, inequality and mathematics performance: The case of South Africa's post-apartheid context. *ZDM, Mathematics Education*, 46(7), 1039–1049.

Graven, M. (2015). Going back in order to go forward – recovery of mathematical foundations for intermediate phase improvement. *Journal of Educational Studies*, 2015(si–1), 1–15.

Graven, M., Venkat, H, Askew, M., Bowie, L., Morrison, S. & Vale, P. (2020a). Grade 3 Mental Starters and Assessments: Print Masters (Creative Commons license) – English.

Graven, M., Venkat, H, Askew, M., Bowie, L., Morrison, S. & Vale, P. (2020b). Grade 3 Mental Starters and Assessments: Teacher Guide (Creative commons license) – English

Graven, M., & Stott, D. (2012). Conceptualising procedural fluency as a spectrum of proficiency. In *Proceedings of the 18th Annual National Congress of the Association for Mathematical Education of South Africa (AMESA)*. 146–156. Potchefstroom: North-West University.

Graven, M., & Venkat, H. (2014). Primary teachers' experiences relating to the administration processes of high-stakes testing: the case of Mathematics Annual National Assessments. *African Journal of Research in Mathematics, Science and Technology Education*, 18(3), 299–310.

Graven, M. & Venkat, H. (2021). Piloting diagnostic assessment for strategic calculation. *Mathematics Education Research Journal*, 33, 23–42

Hedges, L.V. & Schauer, J. (2018). Randomised trials in education in the USA. *Educational Research*, 60:3, 265–275.

Jacobs, V. A., Lamb, L. L. C., & Philipp, R. A. (2010). Professional noticing of children's mathematical thinking. *Journal for Research in Mathematics Education*, 41, 169–202.

Kilpatrick, J., Swafford, J., & Findell, B. (2001). *Adding it up: helping children learn mathematics.* Washington, DC: National Academy Press.

McKay, V. & Spaull, N. (2022). Changing the 'grammar of schooling' in South Africa: The case of the DBE Workbooks. In Spaull, N. & Taylor, S. (Eds) *Early Grade Interventions in South Africa*. Cape Town: Oxford University Press.

Mestry, R. (2017). Principals' perspectives and experiences of their instructional leadership functions to enhance learner achievement in public schools. *Journal of Education*, 69, 257–280.

Metcalfe M., & Witten A. (2019). Taking Change to Scale: Lessons from the Jika iMfundo Campaign for Overcoming Inequalities in Schools. In: Spaull N., Jansen J. (eds) *South African Schooling: The Enigma of Inequality.* Springer, Cham.

Pournara, C. (2020). Grade 9 learners' algebra performance: Comparisons across quintiles, insights from errors and curriculum implications. *South African Journal of Science*, 116 (9/10).

Schollar, E. (2008). *Final report: The primary mathematics research project 2004–2007 – towards evidence-based educational development in South Africa*. Johannesburg, South Africa: Eric Schollar & Associates.

Stott, D., & Graven, M. (2013). Research tensions with the use of timed numeracy fluency assessments as a research tool. *Perspectives in Education*, 31(3), 79–93.

Stott, D., Graven, M., Baart, N., Hebe, G. & Mofu, Z. (2017). After school maths clubs: Investigating learner progression in an expanding intervention model. In T. Penlington & C. Chikiwa (Eds), *Proceedings of the 23rd Annual National Congress of the Association for Mathematics Education of South Africa*, 313–324. Port Elizabeth: AMESA.

Venkat, H. & Askew, M. (2021). Development in South African primary mathematics teacher educators' work with in-service teachers. *International Journal of Science and Mathematics Education* 19 (Suppl 1), 39–58.

Venkat, H., Askew, M., & Morrison, S. (2020). Shape-shifting Davydov's ideas for early number learning in South Africa. *Educational Studies in Mathematics*. Advance online publication.

Venkat, H., & Naidoo, D. (2012). Analyzing coherence for conceptual learning in a Grade 2 numeracy lesson. *Education as Change*, 16(1), 21–33.

Weitz, M., & Venkat, H. (2013). Assessing early number learning: How useful is the Annual National Assessment in Numeracy? *Perspectives in Education*, 31(3), 49–65.

13

Lessons learned and evidence of impact: Formative assessment in an integrated reading and mathematics intervention

ANIL KANJEE & JAYESH BHANA

Abstract

This chapter reports on a seminal intervention programme implemented to support teachers' use of formative assessment to improve learning. The rationale stems from research indicating that while the effective use of formative assessment can result in improved learning gains for all learners, teachers must enhance their pedagogical practices to first identify specific learning needs of their learners before effecting interventions that address these needs.

The intervention was implemented using the ReMAPS professional development model, which is based on a cyclical approach comprising workshops and school-based activities to support teachers to implement and reflect on their experiences as they developed their theoretical knowledge and practical skills. The evaluation was implemented as a randomised control trial by an external agency. Data was obtained from 268 teachers using a reflection exercise, while lesson observations and interviews with learner focus groups were conducted in a sub-sample of schools. Hierarchical linear modelling and chi-square tests were used to analyse the data.

The results revealed that the intervention was successful in supporting teachers, across no-fee and fee-paying schools, to improve their knowledge and practice of several formative assessment

KEYWORDS
formative assessment, randomised control trial, teacher professional development, ReMAPS model

strategies and techniques, attesting to the ReMAPS model for large-scale professional development programmes. However, many teachers still struggled with providing effective feedback and applying peer and self-assessment strategies.

Key lessons indicate that future programmes should provide adequate time for teachers to master all the strategies, while schools should also be supported to adequately plan and prepare before implementing the formative assessment approach. Similar programmes should also be provided for subject advisors to support the professional development needs of teachers. However, additional research is required to determine the extent to which such approaches lead to improved learning gains.

1 Introduction: The challenge of improving learning in SA schools

This chapter reports on a seminal intervention programme intended to help teachers improve their pedagogical practices through the effective use of formative assessment (FA) to identify the learning needs of all learners in their class. The programme contributes to the national initiative for addressing the key challenge of equity and quality that has bedevilled the South African schooling system since the start of the democratic dispensation in 1994. Moreover, given the substantial disparities between schools, the programme also sought to gain a deeper understanding for implementing effective professional-development interventions that support teachers in enhancing their pedagogical practices across both fee and no-fee schools. The focus on FA stems from a plethora of studies indicating that for effective learning to take place, teachers must be able to apply relevant pedagogical practices to determine what learners have learned before implementing any interventions that address their learning gaps. (Furtak, et al. 2008; Harlen 2012; Smith et al. 2005; Stiggins et al. 2006; Wiliam & Thompson 2007). Recent research on the effective implementation of FA approaches in schools not only indicates an improvement in learning gains (Andersson & Palm 2017), but also that these gains were found for all learners, irrespective of their socio-economic backgrounds (Black & Wiliam 1998; OECD/CERI 2008).

In South African schools, however, two key challenges must first be addressed. Firstly, most teachers do not possess the requisite knowledge and skills to effectively use assessment data to identify learning gaps among their learners (Hoadley 2012). Recent results from several studies indicate that South African teachers have limited knowledge and experience in using assessment for improving learning and teaching (Kanjee 2020; Kanjee & Mthembu 2015; Kuze & Shumba 2011; Mkhwanazi et al. 2014). Secondly, guidance and support for teachers are limited, since most district officials are not able to provide the support that teachers require to improve learning and teaching in schools (Nkambule & Amsterdam 2018). Moreover, there is a dearth of information regarding challenges in capacity, cost, and time for implementing large-scale intervention programmes in South Africa that focus on improving teachers' pedagogical practice in schools across the quintiles. In addition, current discourse and practices in schools are dominated by a preoccupation with measurement and testing, manifested by the primary focus on the recording and reporting of learner

scores, with limited use of assessment for enhancing learning (Kanjee & Sayed 2013; Le Cordeur 2014).
Given the specific assessment challenges faced by teachers in South African schools, the application of the FA approach provides a promising alternative for addressing the twin challenges of equity and quality. Within this context, the Formative Assessment Professional Development Programme (FA-PDP) was implemented by the Assessment for Learning Research Niche Area, located at the Tshwane University of Technology, in collaboration with two districts. The primary purpose of the FA-PDP was to enhance teachers' pedagogical practices to effectively use FA approaches for identifying the specific learning needs of all learners in their classroom, as the first step before any interventions can be implemented to address these needs. However, at the request of the district, a parallel programme was implemented for district subject advisors. This chapter focuses specifically on the teachers. The key research questions addressed are: (i) what is the impact of the FA-PDP intervention on teachers' FA knowledge and understanding? and (ii) what is the impact of the FA-PDP intervention on teachers' classroom-assessment practices?

In this chapter, we report on the key lessons learned and their implications for large-scale professional development programmes for teachers in South Africa. The chapter begins with an overview of the FA-PDP that outlines the theory of change, as well as the theoretical, intervention and evaluation frameworks applied. Next, we present the results of the randomised control trial implemented to determine the impact of the intervention on teachers' FA knowledge, understanding, and practices, followed by a discussion of the key lessons learned. The chapter concludes by outlining the key implications for practice as well the next steps for research and intervention.

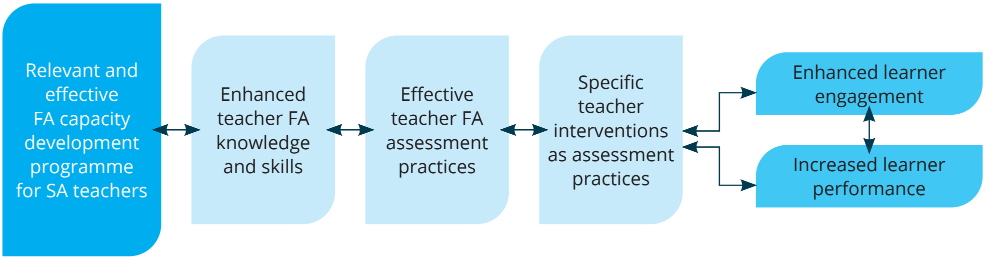
2 Overview of the FA-PDP

In providing an overview of the FA-PDP, we begin by listing the theory of change.

2.1 Theory of change

The theory of change that guided the development and implementation of the FA-PDP (Figure 1) was based on the following: If you implement a formative-assessment-capacity-development programme to support teachers in enhancing their pedagogical practices, and effectively using assessment to identify specific learning needs of all learners, and teachers develop and implement specific interventions to address these learning needs, this will result in enhanced learner engagement and improvement in learner performance.

Figure 1 Theory of change applied to the FA Professional Development Programme



2.2 Theoretical framework

To support teachers' effective use of FA in the classroom, Wiliam and Thompson (2007) identified five key strategies that teachers must apply. The authors based their work on Ramaprasad's (1983, cited in Wiliam 2011) three key processes in learning and teaching: where learners are going, where learners are, and what needs to be done with regards to achieving identified learning goals. Wiliam and Thompson (2007) conceptualised effective FA practices comprising five key strategies for improving learning and teaching: (i) clarifying and sharing learning intentions and success criteria for assessment with learners; (ii) engineering effective classroom discussions, activities, and tasks that elicit evidence of learning; (iii) providing feedback that moves learners forward; (iv) supporting learners to serve as learning resources for each other; and (v) supporting learners to take greater ownership of their own learning. Table 1 shows how the five key FA strategies, three processes, and key agents in the classroom can be linked to provide an integrated framework for FA practice.

Table 1 Application of the key FA strategies

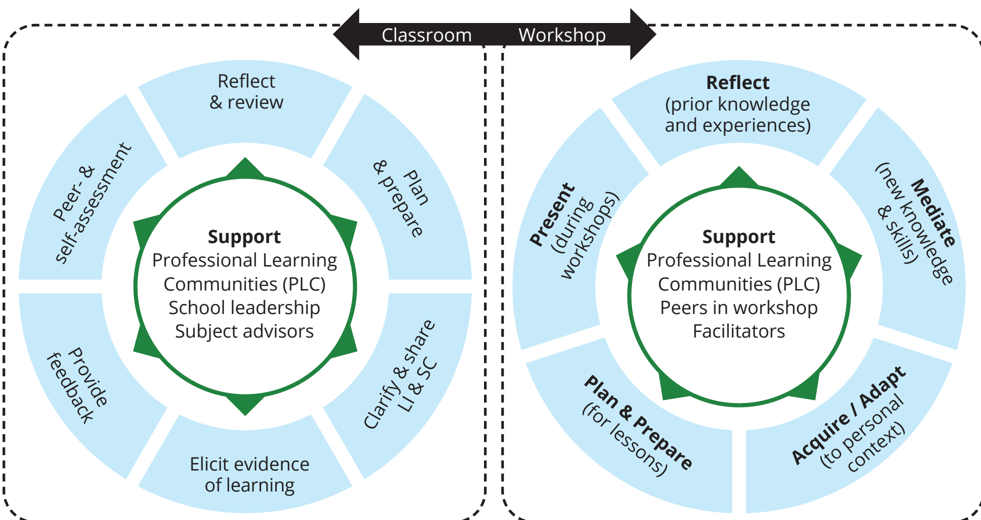
	Where the learner is going	Where the learner is right now	How to get there
Teacher	Clarifying and sharing learning intentions and criteria for success (i)	Engineering effective classroom discussions, activities and tasks that elicit evidence of learning (ii)	Providing feedback that moves learners forward (iii)
Peer	Understanding and sharing learning intentions and criteria for success (i)	Supporting learners to serve as learning resources for each other (iv)	
Learner	Understanding learning intentions and criteria for success (i)	Activating learners as the owners of their own learning (v)	

2.3 Intervention framework

The intervention was based on the ReMAPS model (Figure 2) that was developed for the South African education context (Kanjee & Bhana 2013). This model is based on a cyclical approach where participants collaborate with facilitators to: (i) *reflect* on their current knowledge and skills, (ii) *mediate* new knowledge and understanding during the workshops, (iii) *acquire/adapt* new knowledge and skills to their specific contexts, (iv) *plan and prepare* for implementation and (v) *present/practise* their newly acquired knowledge and skills in the classroom. This cyclical process is underpinned by support from professional learning communities (PLCs), peers, facilitators, and online resources, and is repeated across the series of workshops.

The facilitation, as well as the underlying philosophy advocated, during the workshops is based on reflective teaching and modelling new skills introduced (Korthagen et al. 2006). All workshops comprised presentations, practical applications, video demonstrations, and action planning for application within classrooms. Each workshop was facilitated by experienced members of the Assessment for Learning Research Niche Area. Facilitators also modelled all assessment strategies and techniques that were introduced during the workshops. For example, during the workshop on learning intentions and success criteria, facilitators used learning intentions and success criteria in their introductions while simultaneously providing the theoretical rationale and using specific FA techniques, e.g., name sticks or pair-and-share, to engage teachers during the workshop. Similarly, peer assessment as well as effective feedback strategies were used when teachers were given exercises to review and provide feedback on each other's portfolio exercises.

Figure 2 ReMAPS model applied to the FA-PDP



2.4 Implementation framework

The FA-PDP intervention comprised nine two-hour workshops, as shown in Table 2, which were conducted after school hours at five central venues, and took place over a period of eight months. Each workshop was facilitated by two facilitators to ensure effective support could be provided to all participants. The workshop sessions were provided in clusters of between eight and ten schools and were held at centrally located venues close to the participating schools.

After each workshop, participants were required to implement their newly acquired knowledge and skills in their own classrooms, which included the completion of relevant portfolio activities. They were also required to participate in meetings for the school-based PLC. During the follow-up workshops, participants had to reflect on and share their experiences with colleagues, and adapt any tools and techniques to their learning and teaching contexts.

Table 2 FA workshops offered and content covered

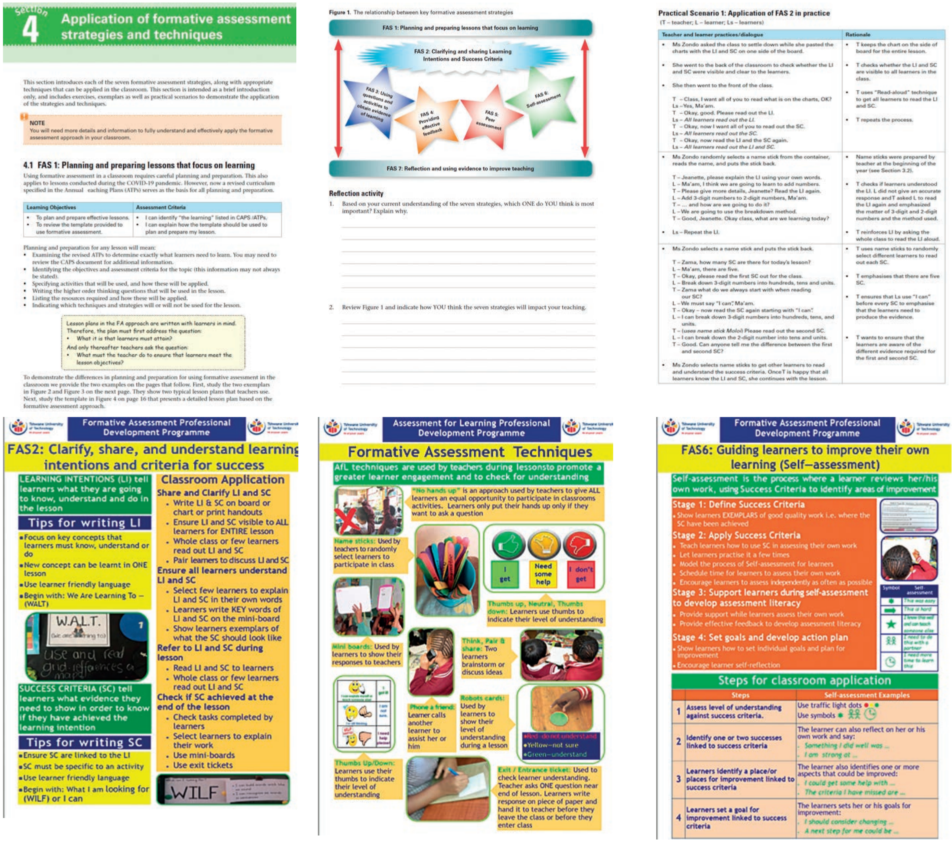
Workshop Number	Content covered
Introduction: Workshop 1	Overview of the programme Establishing teacher learning communities Reflection on assessment practices
School-based activities	
Module 1: Workshops 2, 3 & 4	Introduction to formative assessment Clarifying and sharing learning intentions and criteria for success
School-based activities	
Module 2: Workshop 5	Engineering effective classroom discussions to elicit evidence of learning
School-based activities	
Module 3: Workshops 6 & 7	Providing feedback that moves learners forward
School-based activities	
Module 4: Workshop 8	Assessment by learners of their peers and their own work
School-based activities	
Conclusion & evaluation: Workshop 9	Using assessment to improve my teaching Reflection and evaluation Next steps
Planning for implementation in next school year	

The intervention also included school leaders to enhance the support provided to teachers. Firstly, all Foundation Phase HODs of the selected intervention schools were required to accompany their Grade 2 teachers to participate in the workshops. The inclusion of HODs was vital as they serve as the first-call for support for teachers. Secondly, three workshops were conducted to brief the school management team on the content of the FA-PDP and to propose appropriate systems for supporting teachers to effectively implement FA in their classrooms.

2.5 Learning materials used

Given the absence of any relevant text, learning material in the form of a teacher portfolio, including posters, was provided to all participants. The portfolios were developed to take into account the context of South African schools, as shown in Figure 3, and thus: (i) addressed the reality that English, the language of learning and teaching used for all workshops, was the second language of most teachers; (ii) presented relevant theory on assessment, and included practical exemplars to account for the Foundation Phase curriculum (DBE 2011a; 2011b); (iii) sequenced the content to allow adequate time for teachers to develop and consolidate new theoretical concepts and practical skills learned; (iv) included notes on assessment and education theory, vignettes of classroom experiences, practical tools and templates for classroom use, as well as activities requiring teachers to reflect on their practice, record their lesson preparation and experiences, report on their teacher learning community meetings, and list recommendations or challenges that needed to be shared during workshops; and (v) aligned the course content, activities, and assignments with the teaching plans or programmes of the national curriculum statements.

Figure 3 Example of FA portfolio pages and posters



2.6 Evaluation framework

The evaluation was conducted as a randomised control trial, involving 85 schools, 45 control and 40 treatment, randomly selected across the poverty quintiles in two districts. Qualitative and quantitative data were collected using instruments indicated in Table 3. The success of the intervention was measured based on the following indicators:

1. teachers' FA knowledge and understanding, measured by teachers' responses to

the reflection exercise items;
2. teachers' practices, measured by observations of their FA strategies and techniques during lessons.

3 Methodology

This section describes the sample, the instruments used, as well as the procedures applied for data collection and analysis.

3.1 Sample

Data presented and discussed in this chapter was obtained from a sample of 268 teachers, of which 49% were in the control group and 51% in the treatment group. In addition, 49% were from no-fee schools, and 51% from fee-paying schools. Of the teachers, 57% taught in the Foundation Phase and 43% in the Intermediate Phase, 48% had a foundation phase diploma and 52% a bachelor's degree or higher, while 21% of participants were HODs. In addition, 67% of the participants indicated that their home-language was the same as the language of learning and teaching at the school (i.e., Afrikaans, English, Isizulu, Setswana or Sesotho). Lesson observations were only conducted for a sub-sample of schools (34 schools, 21 intervention and 13 control) while focus-group interviews were conducted with learners after the lesson observations.

3.2 Instruments used

Several instruments were developed to obtain the required data for addressing the key research questions as noted in Table 3.

Table 3 Instruments used and information obtained

Instrument	Target	Information obtained
Reflection Exercise	Teachers – control and intervention groups	Background, teaching experiences, assessment beliefs, practices and knowledge
Observation schedule		Teacher classroom practices, application of FA strategies and techniques; learner participation in classroom activities
Learner focus group interview	Learners – control and intervention groups	Learner knowledge, views and experiences of teachers' FA practices
Review of workbook		Teachers' written feedback practices
Workshop reflection questions	Teachers – intervention group	Teacher experiences of the workshop and/or implementation of the FA strategies and techniques

3.3 Formative assessment reflection exercise (FARE)

The primary purpose of the FA reflection exercise (FARE) was to provide teachers with an opportunity to reflect on their knowledge, beliefs, and practice regarding issues of learning, teaching and FA. The FARE comprised questions regarding teachers' backgrounds, their views of assessment and learning, and 26 multiple-choice questions that probed teachers' knowledge of key FA strategies and techniques. The final results were based on 23 items as three items were removed given low item-total correlations. The Cronbach alpha of 0.729 was considered acceptable given the low-stakes nature of the instrument.

3.4 Observation schedule

The purpose of the observation schedule was to record the FA practices that teachers implemented in their lessons, and included questions on teacher background, teachers' application of the FA strategies, and learner engagement during the lesson.

3.5 Data collection

Data was collected by an independent external agency two weeks after the last workshop was completed. The FARE was administered to all teachers during the final workshop, while teachers in the control group completed the exercise at school after their lesson observations.

3.6 Data analysis

Hierarchical linear modelling (HLM) was used to determine if significant differences existed between and within the subgroups regarding teachers' FA knowledge and understanding. Van der Berg (2008) notes that the use of HLM accounts for some individual effects to be influenced by school-level factors as well as for the random effect that accounts for differences between individual schools. The analyses were performed with the R software (R Core Team 2021) using the lmer, lmerTest and emmeans packages. Chi-square analyses were conducted to determine differences in response rates regarding teachers' use of the FA strategies and techniques as well as learners' responses regarding their classroom experiences and views of their teachers' assessment practices.

4 Results and discussion

In this section the results are presented for teachers' performance on the FARE as well as their classroom practices.

4.1 Teachers' FA knowledge and understanding

A hierarchical linear model was constructed with teachers as level 1 and schools as level 2, with teachers' score on the FARE as the dependent variable. The independent variables were Treat (control vs treatment), Lang (same vs different to the LOLT), Fee (no-fee vs fee-paying), Qual (diploma vs degree), Phase (foundation vs intermediate), and Role (HOD vs teacher). The school regression variable was assumed to be random and independent. The β represent the fixed effects, delta δ is random effects, and ε is the error with ~ $N(0,\sigma^2)$.

$$T_{score} = \beta + \beta_1 * Treat + \beta_2 * Lang + \beta_3 * Fee + \beta_4 * Phase + \beta_5 * Qualif + \beta_6 * Role + \beta_7 * Lang * Treat + \beta_8 * Fee * Treat + \beta_9 * Phase * Treat + \beta_{10} * Qualif * Treat + \beta_{11} * Role * Treat + \delta_{(Sch/Treat)} + \delta_{(Sch/Lang)} + \delta_{(Sch/Fee)} + \delta_{(Sch/Phase)} + \delta_{(Sch/Qualif)} + \delta_{(Sch/Role)} + \delta_{Sch} + \varepsilon$$

As noted in Table 4, a significant effect was only noted for Treatment, Fee and Qualification while no interaction effects between Treatment and other fixed effects were noted.

Table 4 Coefficients for the hierarchical linear model

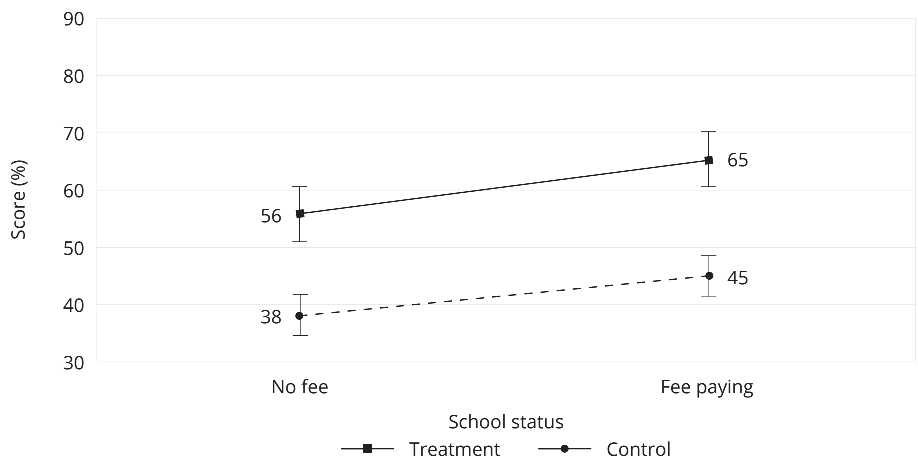
		Estimate	df	t.value	Pr(>|t|)
Intercept	β	38.3693	211.9	7.014	0.0001*
TreatT	β_1	13.5433	199.8	2.111	0.0360*
LangDif	β_2	2.7269	165.9	0.584	0.5603
FeeFee	β_3	6.8988	88.48	2.075	0.0409*
PhaseInt	β_4	−6.0947	129.7	−1.444	0.1512
QualDeg	β_5	3.7915	205.6	1.416	0.1583
RoleTeacher	β_6	−0.7159	208.2	−0.171	0.8645
TreatT:LangDif	β_7	5.2709	124.5	0.860	0.3916
TreatT:Fee	β_8	−2.3069	77.4	−0.469	0.6405
TreatT:PhaseInt	β_9	6.6518	63.3	1.288	0.2024
TreatT:QualDeg	β_{10}	0.7045	200.9	0.205	0.8379
TreatT:HoDTeacher	β_{11}	−1.1069	198.4	−0.230	0.8181
Random effects		**Variance**	**Std Dev.**	**LRT**	**Pr(>Chisq)**
School	δ	17.97	4.23	0.0000	0.9996
School:Lang	δ_1	32.40	5.69	0.9091	0.3403
School:Fee	δ_2	5.04	2.24	0.0000	0.9996
School:Phase	δ_3	7.67	2.77	0.1396	0.7087
School:Qual	δ_4	5.98e−7	0.0004	0.0000	1.0000
School:Role	δ_5	1.14e−7	0.0000	1	0.9997

The contrasts for significant main effects, shown in Table 5, indicate that the treatment effect increased participants' score by 18.1% compared to the control group. Similarly, teachers in fee-paying schools performed 5.75% points higher while teachers with a degree revealed an increase of 4.14%. No differences were found regarding the effects of language of learning and teaching, the status of being a teacher or HOD, nor the effects of teaching across different phases. Moreover, as noted in Figure 4, teachers from no-fee schools who participated in the intervention scored significantly higher than teachers in fee-paying schools.

Table 5 Contrasts for main effects

Contrast	Estimate	df	t-ratio	*p* value
Treat: Treatment vs Control	18.1	124	−5.701	0.0001**
Lang: Same vs Different	5.36	123	−1.679	0.0956
Fee: No fee vs Fee	−5.75	74	−2.285	0.0252*
Phase: Intermediate vs Degree	2.77	59	1.024	0.3100
Qual: Diploma vs Degree	−4.14	68.3	−2.325	0.0230*
Role: HOD vs Teacher	1.27	102	0.511	0.6106

Figure 4 Teacher performance by school fee status and participation in the intervention



The findings indicate that effective professional development programmes, using the ReMAPS model, can lead to significant improvements in teachers' FA knowledge and understanding. More importantly, this improvement was also found among teachers in the low-resourced no-fee schools that primarily serve learners from mainly poor and marginalised backgrounds. However, it is concerning that no difference in performance was detected between teachers and HODs given the critical role of HODs in providing guidance to teachers under their mentorship.

4.2 Teachers' use of FA strategies and techniques

Results from the observations revealed ample evidence regarding teachers' use of the different FA strategies and techniques that were introduced during the workshops. As noted in Table 6, most teachers in the intervention groups applied all the relevant strategies to introduce the learning intentions and success criteria, while no evidence of these strategies was found in lessons observed across the control schools.

Table 6 Teacher application of Learning Intentions and Success Criteria by participation status

FA Strategy	Evidence	Control (%)	Intervention (%)	chi sq	*p* value[1]
Uses Learning Intentions	Seen	0	77	13.109	0.001*
	Not Seen	100	23		
Uses Success Criteria	Seen	0	64	14.267	0.006*
	Not Seen	100	26		
Reminds learners of the LI during lesson	Seen	0	55	12.338	0.015*
	Not Seen	100	27		
Reminds learners of the SC during lesson	Seen	0	55	13.97	0.007*
	Not Seen	100	36		
LI visible throughout the lesson	Seen	0	64	18.250	0.001*
	Not Seen	100	27		
SC visible throughout the lesson	Seen	0	59	19.889	0.001*
	Not Seen	100	32		

Regarding teachers' use of questioning, Table 7 indicates that the random selection of learners was the only strategy which the majority of teachers in the intervention group was seen implementing, and which was significantly different compared to the control group. For the rest of the strategies no significant differences were noted between the intervention and control groups. The results indicate that while most teachers in the intervention groups often used the questioning strategies introduced during the workshops, these same strategies were also applied by teachers in the control group. What this finding alludes to is that these strategies are relatively well-known and applied by most teachers within schools.

Table 7 Teacher use of questionnaire strategies by participation status

FA Strategy	Evidence	Control (%)	Intervention (%)	chi sq	*p* value[2]
Questions are linked to the LI	Often	20	27	4.487	0.106
	Not Seen	80	73		

1 * indicates significance at 0.01 level
2 * indicates significance at 0.01 level

FA Strategy	Evidence	Control (%)	Intervention (%)	chi sq	*p* value[2]
Randomly selects learners	Often	20	80	10.693	0.030*
	Sometimes	0	7		
	Not Seen	80	13		
Learners put hands up when teacher ask questions	Often	80	54	2.479	0.648
	Sometimes	0	33		
	Not Seen	20	13		
Only solicits response from learners with the hands up	Often	40	13	2.332	0.675
	Sometimes	20	27		
	Not Seen	40	60		
Involves more than one learner to respond to question	Often	0	27	4.676	0.322
	Sometimes	20	20		
	Not Seen	80	53		
Questions are targeted to the entire class	Often	60	67	3.728	0.512
	Sometimes	0	13		
	Not Seen	40	20		
Allows some wait-time before learners respond	Often	20	47	3.721	0.514
	Sometimes	40	20		
	Not Seen	40	33		

With regard to teachers' use of the FA techniques, significant differences were noted across all the techniques introduced during the workshops. While the majority of teachers in the intervention group applied these techniques, limited evidence of their use was observed across the control schools. However, lower percentages of teachers were observed applying those techniques that required additional resources that had to be purchased or made, e.g., pass-the-ball technique, robot cards, phone-a-friend.

Table 8 **Teacher application of FA techniques by participation status**

FA Techniques	Evidence	Control (%)	Intervention (%)	chi sq	*p* value[3]
Name sticks	Seen	38	81	8.401	0.015*
	Not Seen	62	19		
T Ball technique	Seen	0	29	7.374	0.025*
	Not Seen	100	71		
Mini boards	Seen	39	62	7.417	0.025*
	Not Seen	61	38		
Robot cards	Seen	23	43	7.037	0.030*
	Not Seen	77	57		
Phone-a-friend	Seen	0	23	11.481	0.003*
	Not Seen	100	77		

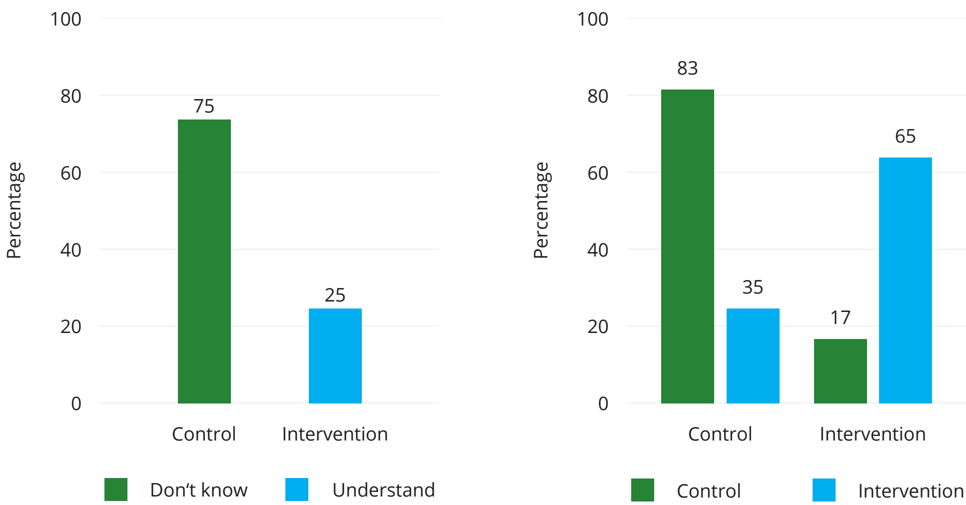
3 * Indicates significance at 0.01

With regard to teacher feedback practices, a review of learners' workbooks revealed no differences in the nature and type of written comments provided by teachers. Across both the intervention and control groups, teachers still demonstrated the predominant use of ticks, crosses, question marks, as well as short comments like 'Good', 'Well done', 'Incomplete'. It seems that many teachers in the intervention group still struggled with providing effective written feedback. Moreover, the use of peer or self-assessment strategies was only observed in one lesson presented by a teacher in the intervention group. However, a key point to acknowledge is that these strategies are only occasionally and selectively applied by teachers, and it may be possible that during a random, single-lesson observation, the specific application of peer- and self-assessment strategies were not used.

4.3 Learner engagement with FA strategies and techniques

Given that the effective use of relevant FA strategies and techniques was intended to support teachers in improving learners' understanding and participation during lessons, the views of learners on their teachers' use of these techniques was considered critical for determining whether teachers applied their new assessment knowledge and skills during lessons. To determine whether teachers presented the learning intentions and success criteria respectively, learners were asked to indicate their understanding of the meaning of the acronym WALT (We Are Learning To) and WILF (What I am Looking For). The chi-square analysis indicated significantly higher percentages of learners in the control group responding 'Don't Know' compared to learners in the intervention groups, x^2 (2, n = 199) = 47.72, p <0.01, regarding their understanding of the term WALT as well as their understanding of the term WILF, x^2 (2, n = 160) = 34.21, p <0.01. These results provide additional evidence that teachers exposed to the intervention were applying some of their newly acquired FA knowledge and skills during lessons.

Figure 5 Learner's understanding of WALT (left) and WILF (right)



Information on learners' exposure to different FA techniques was also obtained during the learner interviews. As reported in Table 9 the chi-square analysis revealed significant differences in four of the six FA techniques reported on by learners. These responses corroborate what was observed during the lessons indicating that teachers who attended the workshops generally applied the FA skills during their lessons. Regarding their use of margin-symbols, responses corroborate teachers' limited use of this feedback practice they were exposed to during the intervention. From this evidence, it can be concluded that learners were also exposed to, and aware of, the specific FA strategies and techniques.

Table 9 Learner responses regarding teachers' use of FA techniques

FA Technique	Learner response	Control (%)	Intervention (%)	chi sq	*p* value[4]
Name sticks	Do not know/use	83	25	60.414	0.000*
	Used sometimes	17	75		
Ball technique	Do not know/use	88	56	20.289	0.000*
	Use at sometimes	12	44		
Mini boards	Do not know/use	79	49	18.494	0.000*
	Used sometimes	21	51		
Robot cards	Do not know/use	92	64	29.460	0.000*
	Used sometimes	8	36		
Phone-a-friend	Do not know/use	82	69	3.297	0.192
	Used sometimes	18	31		
Margin symbols	Do not know/use	58	52	3.253	0.172
	Used sometimes	42	48		

4.4 Learners' interview responses

The interview responses indicate that most learners who were exposed to the FA approach by their teachers responded positively to the use of these strategies and techniques. Many learners reported that they enjoyed the use of name sticks, basketball, and mini-boards as these provided them with more opportunities to participate during lessons. Some of the responses of learners to the question "Which is your best

4 * Indicates significance at 0.01 level

technique?" include: "Name sticks because everyone gets an equal chance to answer"; and "Mini-boards because we can rub the answers easily when we make mistakes".

Regarding the teacher's lesson introduction, learners felt that the use of learning intentions provided them with a better understanding of the lesson purpose. As one learner noted, "it helps us not to forget what we are learning". Regarding success criteria, learners noted that it clarified what they needed to do at the end of the lesson. One learner noted, "it shows us the steps we need to follow," while another learner noted "it shows us different ways to get the answer". One of the most common responses to the reasons why learners approved of these new approaches was aptly summarised by the following response "Now Ma'am gives us a chance to talk in class".

Notwithstanding the positive impact of the intervention on learners, these findings are also concerning given the limited effective written feedback provided to learners. More concerning, however, is the implication that teachers' pedagogical practices have provided limited opportunity for learner participation in classroom activities and/or for effective engagement with the lesson content. It is a serious indictment on the current education system in general, and teachers' pedagogical practices in particular, if learners highlight "the opportunity to participate in class" as one of the most important benefits of the FA approach.

4.5 Teacher responses from course evaluation

Information from teachers' responses obtained from the evaluation of the course revealed that the strategies and techniques were considered to be useful for most teachers. These teachers noted that their participation in the FA-PDP assisted them with planning and presentation of lessons, enhanced their use of teaching strategies which resulted in their learners being more attentive and participating more actively during lessons, while they could easily identify learners who did not understand the content presented. Many teachers also indicated that their learners required more time to familiarise themselves with the new strategies and techniques introduced. In some cases it proved to be difficult for teachers to effectively apply the new strategies during the lessons without causing disruptions. However, some teachers noted that after some time, their learners began to understand and respond to the strategies and techniques. The specific techniques such as name sticks, mini-boards, pair-and-share, and robot cards were regarded as especially useful, as these were easy to introduce and apply and learners responded positively to their application. More importantly, most teachers reported that it allowed more learners to participate in classroom activities while some indicated that it even improved discipline in the classroom. However, many teachers noted that they do not use the strategies and techniques on a daily basis, but only when they felt it necessary.

Several teachers indicated that the application of the FA process would work best if implemented and planned for the beginning of the academic year when it is most constructive to introduce new approaches to their learners. The provision of effective written feedback was regarded as a challenge by most teachers, primarily because of time constraints when marking learners' books, especially in the case of large classes. Some teachers also reported that they found it difficult to provide written comments as their younger learners were still learning how to read. The one area in which most teachers struggled the most pertains to the use of peer- and self-assessment. Specific reasons provided for this include teacher beliefs that this strategy was not useful or applicable to younger learners as they are not able to fully understand what they need to do or effectively engage with their peers. Another reason noted was that this strategy was only covered during the penultimate workshop and thus teachers had limited time to develop their knowledge and skills in its application.

5 Limitations

Two key limitations impacting the FA-PDP need to be noted. These include: (i) the lack of adequate capacity to provide regular support for teachers between workshops; and (ii) restriction of the lesson observations to a sub-sample of the evaluation sample given the limited number of experienced staff who could conduct such observations to determine the extent to which teachers applied any of the FA strategies and techniques.

6 Implications and lessons learned

The findings of the FA-PDP evaluation indicate that teachers' pedagogical practices, across both fee-paying and no-fee schools can be improved through well-designed professional development programmes. However, identifying learners' learning needs is the first step, as teachers still need to develop appropriate strategies and take remedial actions that address these learning needs. Thus, further research is required to demonstrate whether changes in pedagogical practices can lead to improved learning in the South African context. The evaluation nonetheless highlighted several key lessons for implementing effective professional development programmes to improve teachers' pedagogical practices. First, the ReMAPS model as well as the specially designed portfolios proved successful in supporting teachers in developing and effectively using FA strategies and techniques. However, future programmes should also consider: (i) allocating more time for teachers to develop and implement their FA knowledge and skills; (ii) the timing, duration and location of workshops; (iii) exploring the use of school-based workshops as a delivery model. Second, teachers should be provided with additional in-class support between workshops. In this regard, the involvement of subject advisors as well as functioning of PLCs and access to online material could prove critical. Third, there is a need for practical guidelines on how FA strategies and techniques should be implemented during lessons, especially since current policies provide limited information in this respect (Kanjee & Sayed 2013).

7 Conclusion and next steps

The FA-PDP represents an innovative programme intended to support teachers to improve their pedagogical practices to identify and address learning needs of all their learners. The FA-PDP focused on enhancing teachers' use of FA given its impact on pedagogy (Kanjee & Sayed 2013; Le Cordeur 2014), as well as the limited focus within the country on developing teacher capacity and skills to effectively use assessment for improving learning and teaching.

Implemented as a randomised control trial involving teachers in no-fee and fee-paying schools, the evaluation of the FA-PDP sought to determine whether the FA approach could positively impact on teachers' assessment knowledge and classroom practices as well as learner engagement during lessons. The implementation framework, based on the ReMAPS model, was based on a cyclical approach that comprised a series of workshops to facilitate teachers' reflection of their experiences as they developed specific theoretical knowledge and practical skills to improve their pedagogical practice.

The evaluation of the project was conducted by an external agency, with data obtained on teacher-assessment knowledge and classroom practice, and learner views and experiences. Notwithstanding the complexity of implementing large-scale professional development programmes, the results of the evaluation provide sufficient evidence that the use of the FA approach had a positive impact on teachers' pedagogical practices, as well as learners' engagement during lessons, across both fee-paying and no-fee schools. However, there is a need for additional research on the extent to which the improved pedagogical practices lead to improvement in learning and learning outcomes. A critical next step is to develop and implement this model to determine its impact on the functioning of districts, as well as on the improvement of learning outcomes in schools, especially schools that serve learners from poor and marginalised backgrounds.

Acknowledgements

Funding for the programme was obtained from the Zenex Foundation, the Assessment for Learning Research Niche Area, and the National Research Foundation.

References

Andersson, C., & Palm, T. (2017). The impact of formative assessment on student achievement: a study of the effects of changes to classroom practice after a comprehensive professional development programme. *Learning and Instruction*, 49, 92–102.

Black, P. & Wiliam, D. (1998). 'Assessment and classroom learning' *Assessment in Education*. 5(1), 70–74.

Department of Basic Education (DBE), (2011a). National curriculum statement (NCS): Curriculum and policy statement: English Home Language Foundation Phase Grades R–3, Department of Basic Education. Pretoria

Department of Basic Education (DBE), (2011b). National curriculum statement (NCS): Curriculum and policy statement: Mathematics Foundation Phase Grades R–3, Department of Basic Education, Pretoria.

Furtak, E. M., Ruiz-Primo, M. A., Shemwell, J. T., Ayala, C. C., Brandon, P. R., Shavelson, R. J., & Yin, Y. (2008). On the implementing embedded formative assessments and its relation to student learning. *Applied Measurement in Education,* 21, 360–389.

Harlen, W. (2012). On the relationship between assessment for formative and summative purposes. In Gardner, J. (Ed.) *Assessment and Learning*. London: Sage, 87–102.

Hoadley, U. (2012). What do we know about teaching and learning in South African primary schools? *Education as Change*, 16:2, 187–202.

Kanjee, A. (2020). Exploring teachers' use of formative assessment: Prospects and challenges to improve learning for all. *South African Journal of Childhood Education*. 10(1), 1–13.

Kanjee, A. & Mthembu, J. (2015). Assessment literacy of Foundation Phase teachers: An exploratory study. *South African Journal of Childhood Education*, 5(1), 142–168.

Kanjee, A. & Bhana, J. (2013). Formative Assessment Professional Development Programme: Teacher Portfolio. Pretoria: Tshwane University of Technology.

Kanjee, A. & Sayed, Y. (2013). Assessment policy in post-apartheid South Africa: Challenges for improving education quality and learning. *Assessment in Education: Principles, Policy & Practice*, 20(4), 442–469.

Korthagen, F. A. J., Loughran, J. J., & Russell, T. (2006). Developing fundamental principles for teacher education programs and practices. *Teaching and Teacher Education*, 22, 1020–1041.

Kuze, M.W. & Shumba, A. (2011). An investigation into formative assessment practices of teachers in selected schools in Fort Beaufort. *South Africa Journal of Social Science*, 29(2),159–170.

Le Cordeur, M, (2014). Constantly weighing the pig will not make it grow: do teachers teach assessment tests or the curriculum? *Perspectives in Education*, 32(1), 142–157.

Mkhwanazi, H.N., Joubert, I., Phatudi, N.C., & Fraser, W.J. (2014). Teachers' Use of Formative Assessment for the Teaching of Reading Comprehension in Grade 3. *Mediterranean Journal of Social Sciences*. 5(7), 468–475.

Nkambule, G. & Amsterdam, C. (2018). The realities of educator support in a South African school district. *South African Journal of Education*, 38(1): 1–11.

OECD/CERI (2008). Assessment for Learning: The case for formative assessment. Paper presented at the International Conference Learning in the 21st Century: Research, Innovation and Policy. OECD/CERI, Paris; 15-16 May 2008.

R Core Team (2021). *R: A language and environment for statistical computing.* Vienna: R Foundation for Statistical Computing.

Smith, K. A., Sheppard, S. D., Johnson, D. W., & Johnson, R. T. (2005). Pedagogies of engagement: Classroom-based practices. *Journal of Engineering Education,* 94(1), 87–101.

Stiggins, R.J., J.A. Arter, J. Chappuis, & S. Chappuis. (2006). *Classroom assessment for learning: Doing it right—using it well.* Princeton, NJ: Educational Testing Service.

Van der Berg, S. (2008). How effective are poor schools? Poverty and educational outcomes in South Africa. *Studies in Educational Evaluation*, 34(3), 145–154.

Wiliam, D. (2011). *Embedded formative assessment*. Bloomington, IN: Solution Tree Press

Wiliam, D., & Thompson, M. (2007). Integrating assessment with learning: What will it take to make it work? In Dwyer, D. (Ed.), *The Future of Assessment: Shaping Teaching and Learning*, 53–82. New York: Lawrence Erlbaum Associates.

14

The NumberSense programme: A focus of the Shikaya intervention in South Africa

AARNOUT BROMBACHER & NICKY ROBERTS

Abstract

We know that difficulties with learning mathematics stem from early on in the primary school. What to do – at scale – to ensure that children exiting Foundation Phase are ready to thrive mathematically is less well known. This research reports on interventions that show promise in early grade mathematics, noting both their impact as well as the characteristic design features, which are thought to be driving that change. Design-based approaches to creating well-designed and structured learning and teaching materials, which then clearly articulate the necessary conditions for uptake and fidelity of implementation, are therefore helpful. Drawing first on secondary sources we demonstrate promising impacts of the NumberSense Programme in South Africa and Jordan. We then focus explicitly on the Shikaya intervention, exploring five years of Early Grade Mathematics Assessment (EGMA) data ($n = 11{,}516$) to report on shifts in attainment with each cohort. The preliminary analysis over the first three years was promising, but was then interrupted by the Covid pandemic. The design features of the NumberSense workbooks are described. The impact of the Shikaya intervention model ought to be more rigorously evaluated, to establish the extent to which it holds promise for learning improvements at scale.

KEYWORDS
NumberSense, early grade mathematics, EGMA, design-based approach

1 Introduction

Since the first edition of the NumberSense Mathematics Programme workbooks was printed in 2007, more than three million workbooks have been sold. It is estimated that there are currently between 80,000 and 100,000 learners from Grade R to Grade 7, across some 500 schools as well as a range of different intervention programmes, using the workbooks every year in South Africa. The NumberSense Mathematics Programme is implemented in a wide range of different contexts and ways in South Africa. The workbooks have been used as intended in schools across the quintile spectrum. The workbooks have also been used as 'busy' work or homework to supplement the lessons. They have also been used in a wide range of intervention projects that use the materials to support an aspect of the intervention design (typically in quintile 1 no-fee schools).

One of these intervention projects is the Standard Bank-funded Shikaya project implemented in 12 schools across the Western and Eastern Cape since 2016. In the Shikaya project, the NumberSense programme is implemented as the mathematics programme of the school. Learners receive NumberSense workbooks and teachers receive regular classroom-based support by the NumberSense coaches. The research questions addressed in this chapter are:

1. Focusing specifically on the Shikaya intervention, are there preliminary indications of improvements in learner attainment in the Early Grade Mathematics Assessment (EGMA)?
2. From the perspective of the instructional designer, what are the core design features of the NumberSense Programme model, reflecting on the mathematics materials and professional development support?

2 Literature

There are two bodies of literature on which we draw to frame this study on the NumberSense Programme. First, we draw on the literature pertaining to design-based research methods in education. This domain recognises that randomised control trials (RCTs) are costly, and ought to be done only at the point where the most effective intervention design has been trialled and developed. Second, we consider the education literature, drawing particularly on South African examples, which consider systemic change in education. Here we are concerned about the mechanisms which are considered to be promising to produce measurable change in learning outcomes and are scalable. Drawing on the randomised control trials and quasi-experimental designs we establish what is considered a reasonable measure of improvement in learning outcomes (focusing on mathematics in the early grades).

2.1 Design-based research methods

Design-based research (DBR), at times referred to as 'design research' or 'development research' can be traced back to Ann Brown (1992) and the Design-Based Research collective. Dede et al. refer to DBR as offering

> *A best practice stance that has proved useful in complex learning environments, where formative evaluation plays a significant role, and this methodology incorporates both evaluation and empirical analysis and provides multiple entry points for various scholarly endeavors*
>
> (2009, 16).

Importantly, DBR methods are considered appropriate for social environments such as education settings and are contrasted to RCTs. An RCT investigates whether a new teaching strategy is better than a traditional (normal/usual) teaching strategy.

> *To investigate this question one could randomly assign students to the experimental (new teaching strategy) or control condition (traditional strategy), measure performances on pre- and post-tests, and use statistical methods to test the null hypothesis that there is no significant difference between the two conditions. The researchers' hope is that this hypothesis can be rejected so that the new type of intervention (informed by a particular theory) proves to be better*
>
> (Bakker & van Eerde, 2013, 7).

There are however various limitations of RCTs in education, discussed in the literature (see for example, Engeström 2011; Olsen 2004). We highlight just two major problems for RCTs in education. First, an RCT assumes that we know what works and so have a well-defined treatment for our context. This is often not the case. We have few examples of effective interventions that work at scale, on which we can draw (Besharati & Tsotsotso 2015). To improve learning outcomes in a particular setting and subject domain, a new strategy has to be designed before it can be tested. Second, if we know what works, we still do not know why and when it works. It is therefore helpful to start off with a design-based approach where changes over time are tracked and explored. Only when an intervention is stable and showing some promise should an experimental design (using an RCT framework) be conceptualised.

2.2 Improving learner outcomes across a school system

Fleisch (2018) describes "an educational triple cocktail" – structured quality LTSM, teacher training, and school-based coaching – arguing that this type of model for system-wide improvement in learning outcomes (particularly mathematics and language) is showing promise in South Africa. This draws on work conducted in Gauteng Primary Language and Mathematics Strategy (GPLMS) (see Fleisch et al. 2016 and Fleisch 2018). Hazel et al. noted that

> *across a number of recent meta-review and synthesis studies, interventions that target teachers and aim to enhance the quality of instruction, via the introduction of specific teaching methods and/or capacity building, alongside the provision of LTSM, are identified as promising.*
>
> (2019, 52)

Working in India, Banerjee et al. assert that

> *previous randomised studies have shown that addressing children's current learning gaps, rather than following an over-ambitious uniform curriculum, can lead to significant learning gains.*
>
> (2007)

Considering primary level interventions in reading and mathematics, Banerjee (2007) found that interventions that focus on targeting teaching to the current learning levels of students, such as remedial education (Banerjee et al. 2007) or computer-assisted learning, were effective (with effect sizes on mathematics scores being 0.35 standard deviations the first year, and 0.47 the second year). This preliminary work in India has led to a large-scale intervention 'Teaching at the Right Level' (TaRL) focused on mathematics and reading in four provinces in India (Banerjee et al, 2016). Their finding that replicating the early successes at scale was found to be challenging is instructive:

> *providing only materials was insufficient; trained teachers did not adopt the methodology and instead used the textbooks prescribed for the relevant grades, and when volunteers were placed inside schools, they were used by teachers as assistants to implement traditional methods*
>
> (Banerjee et al, 2016, 4).

The subsequent design iterations considered two models: the first (for an environment with good teaching resources), "relied on teachers to implement the programme; however, it also made sure that teachers had a dedicated time in the day devoted to the programme, and were supported from within the government hierarchy" (Banerjee, et al. 2016, 4) and the second (for very poor teaching environments) made use of NGO staff using an out-of-school 'camp' approach.

Drawing across the two sets of literature shows a need to design, test and refine what works in early grade mathematics before rushing to RCTs. There also appears to be promise in interventions, which target the teacher in her classroom, that offer well-designed instructional materials and support for integrating the new pedagogic strategies into her weekly rhythm of engagement. What is promising, is that the Indian example – teaching children at the right level – shifts pedagogy to a more differentiated approach (drawing on formative assessment data rather than age- and grade-level assumptions). Teaching at the right level is intended to enable teachers to address the particular learning level of children (rather than slavishly following an over-ambitious national curriculum). Attaining such shifts at scale requires structured time and management support in addition to appropriate assessment tools and quality LTSM.

3 The Shikaya intervention: origins and implementation

3.1 Design-based origins of NumberSense

The NumberSense Mathematics programme has been developed over 18 years. It has had the benefit of carefully designed LTSM materials in response to practical needs identified in naturalistic classroom settings in quintile 1 schools. It was only by following a substantial period of trial and adaptation that the impact of the NumberSense treatment was then rigorously measured using Early Grade Mathematics Assessments at baseline and after four years of programme implementation.

In 2006, the first author and two colleagues were engaged to provide classroom-based assistance for the teaching of mathematics to the Foundation Phase teachers of the schools supported by the Bitou 10 Foundation. The focus of the support was on equipping teachers with the pedagogical resources to provide differentiated teaching, i.e. to work with the learners in their classes according to their developmental levels in mathematics ('teaching at the right level'). This meant that the teacher arranged the learners in their class into two or three groups according to their developmental level in mathematics. Having arranged the learners in these groups, the teacher then worked with each group (hereafter called the focus group) for approximately one-third of each lesson, aligning the content of the session to the developmental needs of the learners within the group.

The challenge of the approach was neither the grouping of the learners nor the development of the teacher-led learning opportunities for the focus group. Rather, the challenge was ensuring that the learners who were not in the focus group were productively engaged. The term 'productively engaged' in this context means that the task(s) that the learners are working on independently must be (1) meaningful; (2) clear in terms of what is expected of the learner – the learner must know what they have to do; and, (3) tasks that the learners are able to do as they are pitched at the developmental level of the learner. The Bitou 10 Foundation experience highlighted the need for learner materials that are designed specifically to provide independent learning opportunities that reinforce the work of the teacher-led activity with the class and/or focus group and are pitched at the developmental level of the learners.

The NumberSense workbooks and approach were adapted for the Reading and Mathematics Programme (RAMP) in Jordan from 2012 to 2018. The mathematics objective of the programme was to increase the number of early grade students in Jordanian public schools doing mathematics with understanding. The design of the mathematics component of the RAMP was in line with the design of the NumberSense Mathematics Programme. After a baseline EGMA in 2012, teaching and learning materials were developed and piloted in 51 schools in 2013/14. A remedial component was added in 2014/15. In 2015 USAID awarded RTI International the contract to implement the programme at scale from 2015 to 2019. There were 2,500 public schools that participated in the programme, 18,000 teachers were trained, 1,2 million workbooks were printed, and more than 600,000 students were reached. The endline EGMA study reports that

> *statistically significant gains were ultimately found on all mathematics subtasks from baseline to endline. While the gains were relatively small for the more foundational skills (because the baseline scores were quite high), the gains for conceptual skills, which is the focus of RAMP, were larger. Improvements were particularly large in addition and subtraction L2*[1] *(41.9% to 52.1%), as well as in word problems (57.6% to 63.6%). As with reading, overall mathematics gains were similar across G2 and G3*
> (Stern et al. 2020).

Of particular relevance to this chapter is a study by Roberts (2021) which considers the JumpStart implementation of the NumberSense Programme in Ekurhuleni South district of Gauteng. Roberts and Moloi (2021) use four years of cross-sectional Early Grade Mathematics Assessments (EGMA) data (n = 5.724) from treatment and control schools and report a statistically significant difference in mean attainment on the EGMA assessment in the JumpStart schools (effect size of 0.52). Further improvements are evident after three years (effect size of 0.94). The Shikaya implementation of the NumberSense Mathematics Programme was not as costly or tightly implemented as the JumpStart implementation of the NumberSense workbooks. While JumpStart made use of teaching assistants and tablet-based formative assessment processes (see Roberts 2021), the Shikaya intervention relied on teachers to implement (with the support of a coach).

In sum, the NumberSense LTSMs and approach have been carefully designed and researched through design-based methods at a relatively small scale (with 12 no-fee schools), and has subsequently been adopted at national scale in Jordan. The impact evaluation in Jordan shows gains in conceptual skills with large improvements evident for word problems and complex addition and subtraction (Brombacher et al. 2015). The quasi-experiment with a cross sectional design, conducted in Ekurhuleni district in South Africa, where NumberSense Programme materials were used by JumpStart, also shows promising results. It is therefore interesting to explore the Shikaya project, which made use of the NumberSense Programme materials in a different way to JumpStart, and consider whether a RCT ought to be conceptualised.

3.2 Shikaya implementation

The Shikaya NumberSense Mathematics Programme has been implemented in the Western Cape and Eastern Cape since 2016. The programme targets Grades 1 to 3 learners and teachers in 12 urban no-fee schools, although in selected schools, support has also been provided in Grade R and Grades 4 to 7. In a typical year, some 120 teachers and 5,000 learners participate in the programme, 2,500 classroom-based support sessions are conducted and some 13,000 workbooks are supplied. In the majority of schools, the intervention is conducted in English, in two schools in isiXhosa and in another two schools in Afrikaans.

The Shikaya NumberSense Programme aims to improve learner performance in mathematics. That said, the teacher is considered as the change agent and apart

1 L2 refers to Level 2 addition and subtraction problem which include bridging the tens. Eg 27 + 5 =…bridges a ten (exceeding 30, which is the next 10), while 23 + 5 = …stays within the decade.

from supplying learner materials for learners (Number Sense workbooks for each child) the teacher is the focus of the all training – workshops and classroom-based support. NumberSense workbooks are supplied to learners on a needs basis: as learners complete a workbook, they are supplied with the next in the series. In this way, a differentiated approach to teaching is made possible. The children work on learner books which are 'at the right level' for their mathematical development. The children are not expected to doggedly work through a single national curriculum, with the whole class moving at the same pace.

4 Research design

To answer the first research question – *Focusing specifically on the Shikaya intervention, are there preliminary indications of improvements in learner attainment in the EGMA?* – we used EGMA assessment data and analysed it as follows:

Data was obtained from the administration of EGMA assessments in Grades 1,2, and 3 from 2016 to 2021 in Shikaya schools. It must be noted that because of Covid-19-related disturbances on the study, there is no data for the year 2020. Primary data-cleaning and manipulation was done in Microsoft Excel, whilst secondary data-cleaning and manipulation, together with all analysis was conducted in R (a statistical software programme). The study defined EGMA raw scores by summing up all the individual marks for the different assessments the students had taken. Secondary data cleaning, included taking all the relevant variables (Grade, Year, Raw Scores and School) into one data set. The number of learners writing the EGMA in each grade per year is shown in Table 1.

Table 1 Learners assessed using the EGMA from 2016 to 2021 in Shikaya schools[2]

Cohorts n	Baseline (first data point)	2016	2017 Year 1	2018 Year 2	2019 Year 3	2020 Year 4	2021 Year 5
Grade 1	*n* = 960	*n* = 960	*n* = 1,247	*n* = 771	*n* = 990	No data	No data[3]
Grade 2	*n* = 980	*n* = 980	*n* = 1,191	*n* = 827	*n* = 1,024	No data	*n* = 211
Grade 3	*n* = 1074	No data	*n* = 1,074	*n* = 849	*n* = 981	No data	*n* = 211

2 Note that the colour coding follows a cohort of particular learners: learners who start school in Grade 1 in 2016, are followed into Grade 2 in 2017 and then to Grade 3 in 2018.

3 No data was collected in from Grade 1 learners in 2021. School were operating on a rotation basis and access to all classes was difficult.

For the baseline to 2019, all learners in eight schools were assessed. In four of the schools, 15 learners per class were randomly selected for assessment. The data set in 2021 is diminished as learner sampling (15 learners per class) was conducted in all schools.

The baseline group was defined as 2016 Grade 1, 2016 Grade 2, and 2017 Grade 3. The baseline group was chosen as an internal comparison group as there was no available control group. The baseline group is from the same schools, same teachers, and hence same culture of teaching mathematics. The chosen baseline was of the same context and as close to the start of the intervention as possible. It is assumed that by taking the learner performance in each grade-level at the start of the intervention this reflects the usual or normal performance of the Shikaya schools (prior to the NumberSense intervention being embedded into the schools).

To analyse the EGMA data we defined and colour-coded learner cohorts:

- Shikaya (baseline): 2016 G1, 2016 G2, 2017 G3 (this is shown as black and is the first data point taken for each grade)
- Shikaya year 1: G1 2017, G2 2018, G3 2019 (medium)
- Shikaya year 2: G1 2018, G2 2019, G3 2020 (darker)
- Shikaya year 3: G1 2019, G2 2020, G3 2021 (darkest)

We first calculated the mean and standard deviation for EGMA raw scores in each year and grade-level. To determine whether the Shikaya schools were improving their EGMA performance in each year of intervention, we used a two-sample t-test and 5% significance to determine whether the observed differences between the mean raw scores were statistically significant or not. We performed the following t-tests, comparing:

- Grade 1 year 1 (2017), year 2 (2018), year 3 (2019) to the Grade 1 baseline (2016),
- Grade 2 year 1 (2017), year 2 (2018), year 3 (2019) to the Grade 2 baseline (2016),
- Grade 3 year 2 (2018), year 3 (2019) to the Grade 3 baseline (2017)

Where there were significant differences we calculated Cohen's D and Hedges G to get a preliminary indication of the scale of improvement.

To answer the second question – *From the perspective of the instructional designer, what are core design features of the Shikaya model*? – we drew on the knowledge and experience of the first author who was the primary designer of the materials and the professional development offering. The second author checked the coherence and sense of these descriptions considering the materials themselves, making judgements on whether the claims made were in fact evident in the NumberSense Programme materials and Shikaya project reports.

5 Analysis and findings

5.1 Focusing specifically on the Shikaya intervention, are there improvements in learner attainment in the EGMA?

The preliminary exploration of the EGMA data suggests that there are improvements in learner attainment in the EGMA with engagement with the NumberSense Programme in the Shikaya schools. We cannot, at this point, claim causality – that the changes seen are solely a result of the Shikaya intervention – but the improvements at least suggest that further study relating to impact is warranted.

Table 2 Shikaya EGMA data from 2016 to 2021 by Grade

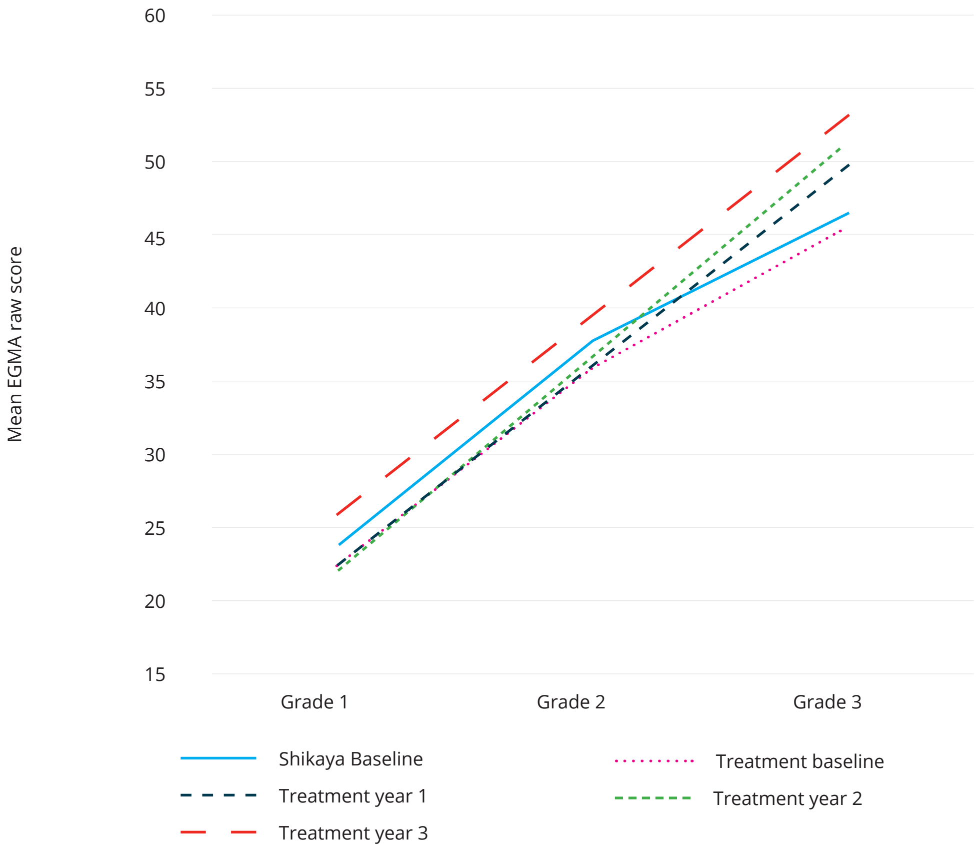
Cohorts *n* Mean (SD)[4]	Base-line	2016	2017	2018	2019	2020	2021
Grade 1	*n* = 960 23.9 (10.9)	*n* = 960 23.9 (10.9)	*n* = 1247 22.5* (9.76)	*n* =771 22.1* (9.23)	*n* = 990 26* (10.1)	No data	No data
Grade 2	*n* = 980 37.7 (12.8)	*n* =980 37.7 (12.8)	*n* = 1191 36.0* (12)	*n* = 827 36.5* (12.4)	*n* =1024 37.3 (13.4)	No data Estimate: $\frac{(42.6 + 37.3)}{2}$	*n* = 211 42.6* (11.5)
Grade 3	*n* = 1074 46.4 (12)	No data	*n* = 1074 46.4 (12)	*n* = 849 45.6 (12.8)	*n* = 981 49.5* (12.4)	No data Estimate: $\frac{(53.1 + 49.5)}{2}$	*n* = 211 53.1* (13.1)

**Significant difference compared to baseline (p<0.05)*

The learner attainment clearly distinguished learners by grade. Learners in higher grades performed better than those in lower grades. This is expected as older students tend to do better than their younger counterparts. In addition, as the Shikaya programme was embedded into the schools, performance at each grade-level improved over time. The mean EGMA results by cohort increased with each year of NumberSense intervention.

4 Each cell reflects three values: *n* (the number of learners tested); the mean result on the EGMA, and the standard deviation (sd) from the mean (shown in brackets).

Figure 1 Shikaya EGMA attainment by grade over time



The treatment year 1 learners entered the schools in Grade 1 slightly weaker than the baseline cohort. This weaker performance remained evident in Grade 2, but by Grade 3 they were performing better than the baseline. The treatment year 3 cohort entered the school stronger than the baseline and sustained this better attainment when they reached Grade 2 and Grade 3.

We expect that Grade 1 learners drawn from the same communities that enter the same group of schools will perform similarly over time. This was generally the case as the difference in mean result for the year 1 and year 2 Grade 1 learners, compared to the Grade 1 learners at the baseline was negligible. The year 3 Grade 1 group was significantly better than previous groups, but this difference was small (D = 0.2). The same trend is evident for the Grade 2 learners. They all perform similarly to the baseline, but by year 5 (2021), after Covid-19, we see a higher attainment that is significant, with a small effect size (D = 0.39). By Grade 3 our exploratory analysis suggests greater impact of the NumberSense Programme. In year 2 of the intervention there was no significant difference in the Grade 3 attainment. By year 3, there was a small improvement (D = 0.25) and by year 5 (after Covid-19) the mean result was significantly higher than the baseline with a medium effect size (D = 0.54). Cohen's D and Hedges G are almost equal for all the comparisons (giving the same measure of the difference in means). The absolute values of the effect sizes increase each year from 2017 with the lowest to 2021 having the highest. This trend is noticed in all years of the intervention.

Our preliminary exploration of EGMA attainment from the baseline, over time, suggests that the design being used is worth further research to establish impact. The way of working is therefore of relevance, and we turn now to describing its design features.

5.2 From the perspective of the instructional designer, what are the core design features of the Shikaya model?

In addition to the structured routines that the programme provides for teaching and learning mathematics, the appeal and success of the programme lies in the underlying philosophy. This philosophy assumes that all learners are inquisitive, sense-making, problem-solving individuals, capable of learning mathematics from deliberately and thoughtfully structured activities designed to reveal the mathematics and mathematical heuristics, which learners need to develop.

Consistent with design-based approach, the Shikaya programme presented an opportunity to test the materials and implementation philosophy in 'real' classrooms, i.e. to develop a better understanding of how realistic the implementation methodology is, in particular in low socio-economic schools and classrooms (average #learners ≈ 38).

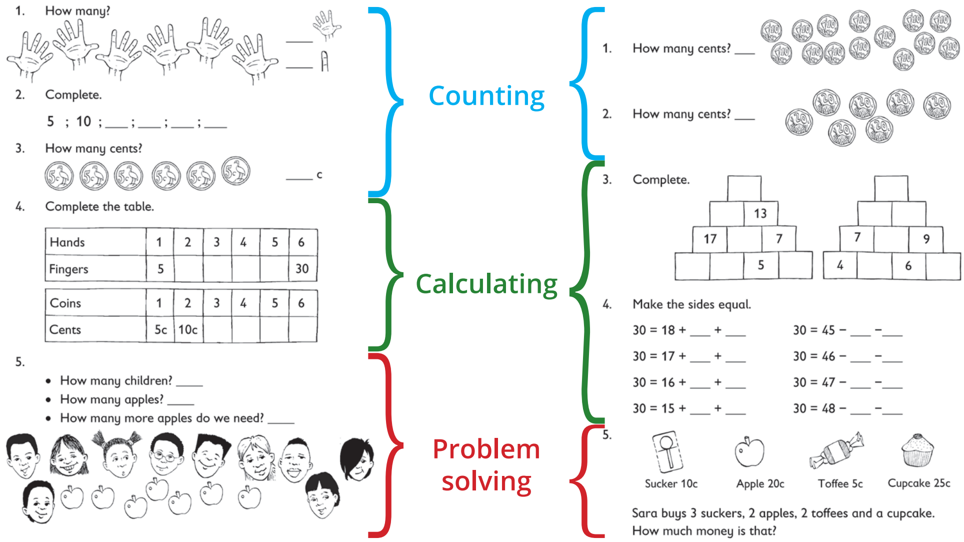
Next, we focus on the instructional design features of the NumberSense learning materials, then on the professional development support offered to teachers.

5.2.1 Key design features of the NumberSense workbooks

The NumberSense mathematics materials costed at R30 per book for the programme (R56 per book retail). An average learner works through three books per year (R90 per year at cost). The workbooks have a simple design that supports learners becoming familiar with the page layout and activity devices on a page. Figure 2 illustrates two typical workbook pages.[5] Learners' familiarity with the pages, in turn, develops learner confidence in completing the activities. The NumberSense materials are expressly designed to enable learners to independently complete a page a day for each school day of the year.

5 While a child may not notice the shifts from counting to calculating to problem solving, this repetition of the three aspects on each page offers structured support of appropriate activity types to teachers.

Figure 2 Sample pages from the NumberSense



The NumberSense Mathematics Programme was designed to be much more than a set of workbooks. It was designed as a highly-structured approach to teaching mathematics in the early years. It was primarily designed to support teachers in providing a robust, thoughtfully sequenced mathematics learning programme. Repeating the same page structure allows teachers to develop a ritual of working on a page a day, while ensuring that all three tasks – counting, calculating, and problem-solving have been given some attention.

The nature of the tasks and the underlying philosophy of the programme is to develop mathematical proficiency. That is, knowledge with understanding, which can be applied in unfamiliar situations, all the while reasoning about why and how the knowledge has been applied to make sense of the situation and/or solve the problem (Kilpatrick et al. 2001).

An important design feature of the materials is the expectation that learners should be able to work with the materials independently of the teacher. To achieve this objective, the pages all have a similar format using a limited number of different activity devices. The activity devices (flow diagrams; number chains; tables; pyramids and so on) used in the early workbooks are the same activity devices used in the later workbooks. All that changes are the number range and the complexity of the mathematics as learners progress through the workbooks.

Another important design feature of the materials is that while broadly aligned to a typical mathematics curriculum for the early grade, the mathematics is not presented in chapters/sections according to curriculum topics.[6] This is deliberate. The mathematics of the early years in highly interrelated: it is possible to solve almost all subtraction problems using addition, all division problems using either

6 See Moloi, Roberts and Thomo (this volume) for more detail on how the NumberSense books have been integrated into mathematics lessons.

repeated addition or repeated subtraction, and all multiplication problems using repeated addition. The materials are designed to support learners in developing an understanding for and appreciation of this interrelatedness.
The materials are also not so much guided by the artificial grade-level number ranges of a typical curriculum, but rather by the application of knowledge with understanding. If a learner knows that 5 + 2 = 7, they can apply this knowledge to calculate

$25 + 2 = \square$;
$85 + 2 = \square$;
$165 + 2 = \square$; and
even $500 + 200 = \square$; and
$5{,}000 + 2{,}000 = \square$.

To take this one step further, we know that the approach is working when a learner explains that 63 + 4 = 67, because 3 + 4 = 7. The materials also, and deliberately, present

$5 + 2 = 7$ as
$5 + 2 = \square$;
$5 + \square = 7$;
$\square + 2 = 7$, and also as:
$7 = \square + 2$;
$7 = 5 + \square$;
$\square = 5 + 2$ etc.

to facilitate the development of a robust understanding of the equals sign as denoting equivalence, as opposed to a symbol that precedes 'the answer.'

Key to developing learner confidence with mathematics is the spiral nature of the materials: concepts are continuously revisited with a gradual increase in cognitive demand supported by the revisiting of the concept(s) in less cognitively demanding situations.

In the Grades 1 to 3 workbooks, a page (and by extension that teacher-led activity) typically has three elements: a counting activity, a manipulating number (mental arithmetic) activity, and a problem to be solved. This is shown in Figure 2 which shows these three elements on typical pages. The amount of time spent on each of these three elements varies according to grade. At the start of Grade 1, most of the teacher-led activity is spent on counting and the remainder of the time on solving a problem. Over the years, the amount of time spent on counting is reduced, and manipulating numbers and solving problems take up more of the time.

The purpose of the counting activity is to support the development of learners' numerosity (their sense of muchness) – learners need to understand that 500 is a lot more than five and simply reading the numbers does not convey this. Furthermore, as the quantity being counted increases, learners become aware of the need for increasing efficiency – counting in groups, which lays the foundation for what will one day be multiplication.

Word problems in the NumberSense programme have two purposes. First and foremost, they present situations that learners are able to make sense of and solve using grade- and age-level appropriate strategies (from physical modelling, to using drawings, to primitive and eventually sophisticated number strategies). Second, solving the problem reveals the mathematics that we want learners to develop – the problems provoke an organic reaction, which is the mathematics that we want to reveal. If we want to introduce the concepts of addition and subtraction we use change, combine, and compare problems; for division we use sharing and grouping problems; and to introduce the fraction concept we use sharing problems with remainders that can easily be partitioned and so on.

Manipulating numbers (mental arithmetics) is used to develop a critical mass of number facts (the 'sight words' of mathematics).

A page in a NumberSense workbook represents a lesson (a day). The activities of the page are almost always interrelated with the counting and manipulating number activities linked by a common theme. One of the roles of the teacher is to assist learners in reflecting on the activities on a page so that they become more aware of the patterns and relationships on the page than they might do on their own.

5.2.2 Key design features of the NumberSense professional development offering

The NumberSense Programme workbooks are also intended to support teachers and teacher development. First, the pages of the learner workbooks are, in effect, the teacher's lesson plan. The lesson begins with a teacher-led activity that sets the learner up to independently complete the workbook page as a consolidation of the teacher-led activity. Second, having teachers work through the materials to prepare for their lessons supports the development of a richer, more robust understanding of the mathematics that they are teaching.

Workshops for teachers are typically presented with at least two workshops per school per year and costed at R1,500 per workshop (for the Shikaya intervention, funded by Standard bank). All teachers receive classroom-based coaching.

All but two of the coaches have at a minimum a university degree with a post-graduate certificate in education. The other two coaches have a college teaching diploma. In all instances the coaches were full-time salaried employees of Brombacher and Associates. The time allocated to the Shikaya programme for the degreed-coaches varied between 20% and 50%. For both coaches with college diplomas, they were engaged on the programme on a full-time basis. Each coach is responsible for about 17 teachers.[7]

The Shikaya programme presented the team with an action-research setting in which to trial, test, and improve the programme design in a South African setting. One example of the lessons from the programme is the development of the developmental trajectory for teachers implementing a new programme, which follows: active resistance; reluctant engagement; and spontaneous engagement. As we are aware that effecting change at school through a teacher-led intervention is difficult, a metric for monitoring and reporting on teacher behaviour was developed for the coaches. The teacher-engagement phase is determined through an analysis of the classroom-visit report completed by the coach at the conclusion of each lesson, as well as the quarterly page-rate audit (the number of pages completed by learners in a week) conducted by the coaches for the learners in each teacher's class.

7 Degreed coaches tend to have more experience and knowledge of the mathematics and its teaching. They are more expensive resources who then support the less experienced coaches who have diplomas.

Table 3 Phases of teacher engagement, teacher behaviour and the role of the coach

Phase	Teacher behaviour	Coach's role
Active resistance	The teacher avoids using the routines and/or actively decides not to do so. The teacher finds reasons not to use the routines.	To be (frequently) present and supportive, to remove barriers to using the routines and, if necessary, demonstrate the routines for the teacher.
Reluctant engagement	The teacher uses the routine but still lacks enthusiasm.	Encourage the teacher, highlight aspects of the implementation that are going well and draw the teacher's attention to the positive responses of the students to the routines.
Spontaneous engagement	There is evidence that the teacher uses the routine even when the coach is not visiting.	Facilitate reflection with the teacher on the value of the routine and support the teacher as he/she starts to explore variations of and modifications to the routine.

Teachers receive classroom-based support according to the phases of their engagement with the programme. The three phases including teacher behaviour and the role of the coach are described in Table 3. The frequency of classroom-based support by the NumberSense coach, when teachers are at the active resistance stage, is once or twice per week. Teachers who are at the reluctant engagement stage receive coaching bi-weekly and teachers at the spontaneous engagement phase receive monthly support. On average, a teacher receives 17 lessons per year in which they are supported by a coach. The cost of a coach supporting a lesson is estimated at R200 per lesson.

6 Conclusion and way forward

The Shikaya implementation of the NumberSense Programme, which relies on teachers to implement the programme with the support of a numeracy coach, shows preliminary indications of positive learning gains. With well-designed learning materials, teacher training, and a coach, it is possible to support teachers to teach at the right level. The key design features of the NumberSense Programme learning materials are summarised as being:

1. Provide simple structured routines to support teaching and learning mathematics.
2. Assume that all learners are inquisitive, sense-making, problem-solving individuals, capable of learning mathematics from deliberately and thoughtfully structured activities.
3. Use a simple design that supports learners becoming familiar with the page layout and activity devices on a page.
4. Design learner materials to independently complete a page a day for each school day of the year
5. Design learner materials to also support teachers and teacher development.

6. Expect that learners should be able to work with the materials independently of the teacher, so ensure all pages all have a similar format using a limited number of different activity devices.
7. Do not design the pace of mathematics in learner workbooks to rigidly following the grade level of the curriculum and topic sequence of the recommended teaching programme, but focus, rather, on a robust, research-based developmental trajectory.
8. Do not use the artificial grade-level number ranges of a typical curriculum, but rather increase the difficulty by the application of knowledge with understanding, by, for example, using structural relations (families of related number sentences) and varying the position of the unknown to develop meaningful use of the equal sign.
9. Build learner confidence with mathematics by using the spiral nature of the materials, so that concepts are continuously revisited with a gradual increase in cognitive demand supported by the revisiting of the concepts in less cognitively demanding situations.

In Jordan the mathematics component of the RAMP intervention, which was modelled on the NumberSense Mathematics Programme materials and methodology, showed improved learning outcomes, at a national scale. Stern et al. (2020) report significant gains as a result of the intervention in Jordan. The gains are similar in both Grade 2 and Grade 3, with the greatest improvements evident in addition and subtraction level 2 (41.9% to 52.1%), as well as in word problems (57.6% to 63.6%). In South Africa, Roberts and Moloi (2021) report a statistically significant difference in mean attainment on the EGMA assessment in the JumpStart schools, which make use of NumberSense workbooks (effect size of D = 0.52) and with further improvements evident after three years (effect size of D = 0.94). Applying design-based methods, this paper presents positive learning improvements from preliminary analysis of the Shikaya EGMA data and documenting the key design features of the model. The preliminary analysis suggests that it is now appropriate to conduct a more rigorous randomised control trial, to evaluate the NumberSense Programme materials and professional development model at a larger scale.

References

Bakker, A., & Van Eerde, H.A.A. (2013). An introduction to design based research with an example from statistics education. In BiknerAhsbahs, A. Knipping, C. & Presmeg, N (Eds.) *Doing qualitative research: Methodology and methods in mathematics education*. New York: Springer.

Banerjee, A., Banerji, R., Berry, J, Duflo, E., Kannan, H., Mukherji, S., Shotland, M. & Walton, M. (2016). Mainstreaming an effective intervention: Evidence from randomized evaluations or "teaching at the right level" in India, National Bureau of Economic Research, Working Paper 22746 http://www.nber.org/papers/w22746

Banerjee, A.V., Cole, S., Duflo, E., & L. Linden. (2007). Remedying Education: Evidence from Two Randomized Experiments in India. *The Quarterly Journal of Economics*, 122(3),1235– 1264.

Besharati, N. A., & Tsotsotso, K. (2015). *In search for the education panacea: A systematic review and comparative meta-analysis of interventions to improve learner achievement in South Africa*. Johannesburg: University of the Witwatersrand.

Brombacher, A., Stern, J., Nordstrum, L., Cummiskey, C., & Mulcahy-Dunn, A. (2015). *Education Data for Decision Making (EdData II): National Early Grade Literacy and Numeracy Survey–Jordan. Intervention impact analysis report*. Available at: https://jordankmportal.com/resources/jordan-intervention-report-english (Accessed 14 November 2022).

Dede, C. Jass Ketelhut, D. Whitehouse, P. Breit, L. & McCloskey, E. M. (2009). A Research Agenda for Online Teacher Professional Development. *Journal of Teacher Education*.60(1), 8–19.

Engeström, Y. (2011). From design experiments to formative interventions. Theory and Psychology, 21 (5), 598–628.

Fleisch, B. (2018). *The education triple cocktail: System-wide instructional reform in South Africa*. Cape Town: University of Cape Town Press.

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

Hazell, E., Spencer-Smith, G., & Roberts, N. (2019). Improving Grade R mathematics teaching in South Africa: Evidence from an impact evaluation of a province-wide intervention. *Journal of Education*, 76, 48–68.

Kilpatrick, J., Swafford, J., & Findell, B. (Eds) (2001). *Adding it up: helping children learn mathematics*. Washington, DC :National Academy Press

Olsen, D. R. (2004). The triumph of hope over experience in the search for "what works": A response to Slavin. *Educational Researcher*, 33(1), 24–26.

Roberts N. (2021). Assessing early grade mathematics learner outcomes using m-learning. In: Auer M.E., Tsiatsos T. (Eds) *Teaching Assistant : Intern of Things, Infrastructures and Mobile Applications*. IMCL 2019. Advances in Intelligent Systems and Computing, 1192. Springer, Cham. pp. 136-146.

Roberts, N. & Moloi, Q. (2021). JumpStart Programme Impact on the Mathematics Learning Outcomes: A cross sectional quasi-experiment using the Early Grade Mathematics Assessment. In Qhobela, M., Ntsohi, M.M.E., & Mohafa, L.G. (Eds). *Book of Proceedings of the 29th Annual Conference of the Southern African Association for Research in Mathematics, Science and Technology Education*. Lesotho: SAARMSTE. pp. 122-138

Stern, J., Abdelgawad, A., Fayaud, P. & Al-Jundi, R. (2020). *Early Grade Reading and Mathematics Initiative Endline Survey Report*. RTI International.

15

Using NumberSense workbooks and formative assessment to improve learning outcomes in early grade mathematics

QETELO MOLOI, NICKY ROBERTS & JABU THOMO

Abstract

JumpStart is one of a few examples of South African early grade mathematics interventions that have shown remarkable gains in learning outcomes in the Early Grade Mathematics Assessment (EGMA). Four years of EGMA data, administered annually from 2016-2019 from Grades 1 to 3 (n = 1,865) were analysed, focusing specifically on the simple addition and subtraction testlet. Using the JumpStart baseline data – as a measure of 'likely attainment' in urban no-fee schools – the average child correctly answered two, seven, and 11 simple addition or subtraction calculations in two minutes, in Grades 1, 2, and 3 respectively. Following the JumpStart intervention, and as a measure of 'aspirational attainment levels', the average child correctly answers three, 12, and 16 simple addition or subtraction calculation in two minutes, in Grades 1, 2, and 3 respectively. JumpStart programme participants were interviewed to reflect on what they considered to be the key mechanisms for improving early grade mathematics outcomes. The formative assessment mobile application 'Numeracy tracker' was reviewed. JumpStart worked with a district, applying five key mechanisms, which are considered to be critical to the JumpStart design: (1) teachers and teacher training; (2) NumberSense learner workbooks; (3) individual attention from teaching assistants or

KEYWORDS
EGMA, formative assessment, Number Sense, mathematics, early grades

interns who each have a tablet; (4) coherent mathematics pedagogy; and (5) real-time monitoring of learning (formative assessment). An important outcome of this study is that differentiated pacing, real-time formative assessment feedback, coupled with weekly attention, are mechanisms which are worth trailing in other contexts.

1 Introduction

South African educators and researchers have been aware of persistent under-performance at the primary school level, particularly in the foundational skills of numeracy and literacy (SACMEQ 2017). Not only has performance been low, but study after study showed that the country was dealing with an extremely unequal society mirrored in what Van der Berg (2008) referred to as a bi-modal system where, on the one hand, there is a minority of schools where achievement was relatively high and, on the other hand, there is a majority (approximately 80%) of schools where performance is extremely low. In the last ten years (2010-2020) there have been several interventions, both by the government or government collaborating with the private sector, or by independent organisations with an interest in education, to try and turn the situation around.

In the case of early grade mathematics, the interventions have varied in the approaches that they implemented to trigger change: the level of the education system that they address, the functionaries that they address, and the nature and amount of the dosage that they administer.

There is an emerging body of literature that shows that interventions that leverage on pedagogical practices inside the classrooms have the greatest potential to impact on learners' performance (Black & Wiliam 1998; Desimone & Pak 2017). While positing that the use of strategies of formative assessment can lead to improved learner performance, Kanjee (2020) reported that only a minority of teachers in quintiles 1-5 were able to use appropriate formative assessment strategies in the classroom.

This chapter provides an account of the process and outcomes of the evaluation of the impact of an intervention, with a particular focus on NumberSense, by JumpStart in township schools in Ekurhuleni South, an education district in the Gauteng Province. Roberts (2019) frames the JumpStart intervention as an example of an m-learning intervention where tablets are used by Teaching Assistants (TAs) to monitor and respond to the mathematics learning of Foundation Phase children. The cross-sectional study reported in Roberts (2019) makes use of the Early Grade Mathematics Assessment (EGMA) to demonstrate that the JumpStart programme had a significant impact on the average school in the intervention.

> *Every school has a statistically significant increase in their respective EGMA means (compared to their own baseline, and to the comparison schools). Overall the JumpStart programme is having a significant impact on all the schools. Improvements in learning outcomes, as measured by the EGMA are evident after one year of intervention, and appear to be strengthened after two years of intervention.*
>
> (2019, 145)

In an extension of the Roberts (2019) study, Roberts and Moloi (2021) draw on additional data, using four years of cross-sectional EGMA data (n = 5,724) from treatment and control primary schools in the same district in Gauteng. Early Grade Mathematics Assessments were administered to a random selection of learners in both school groups over a period of four years. Roberts and Moloi (2021) report a statistically significant difference in mean attainment on the EGMA assessment in the JumpStart schools (effect size of 0.52) and with further improvements evident after three years (effect size of 0.94), estimated as Cohen's D co-efficient.

In addition the Roberts and Moloi (2021) study offered benchmarks for what can be expected for EGMA attainment (raw score) in each grade level. However, knowing that children in mainstream schools in this Gauteng district obtain an EGMA raw score of 22, 40, and 50 at Grades 1, 2, and 3 respectively does not offer any meaningful insight in what they are able to do mathematically. Roberts and Moloi (2021) also report the shifts in attainment as effect size (Cohen's D) where the units are in standard deviations. While these benchmark figures may well be helpful to researchers using the EGMA, we do not think they are meaningful or useful to teachers, project implementers, or subject advisors.

Therefore, this chapter sets out to do two things. First it provides some evidence of the improvements in mathematics learner outcomes in the JumpStart schools – focusing on simple addition and subtraction fluency. Second, it unpacks the nature of the JumpStart intervention in more detail, attempting to reveal what is 'inside the black box' of the JumpStart intervention design.

2 Literature review

There is growing literature from impact evaluations that suggests that education interventions that are targeted at the district level are most likely to result in the intended outcomes mainly because 'districts'[1] exercise direct jurisdiction over clusters of schools (Khosa 2014; Mouton et al. 2013). Interventions at the district level have been a common practice presumably because districts, as the closest support structures to the schools, have the potential to sustain the impact of interventions. However, literature on the effect sizes of various interventions in South Africa and elsewhere have thrown up a variety of results with traceable evidence that suggests that interventions with the largest effect sizes are those that impact what happens inside the 'black box', that being the classroom (Black & Wiliam 2009).

In South Africa there have been numerous interventions to improve teaching and learning in early grade mathematics between 2010 and 2020. The interventions varied both in terms of the inputs that were made and the instruments that were used to measure impact although the degree of the impact varied in magnitude. Roberts and Moloi (2021) offer a tabulation of average effect sizes (Cohen's D), drawing on global literature and empirical studies in South Africa. The impact ranges from an effect size of 0.94 for the JumpStart intervention in Grades 2 to 3, after three academic years; to 0.01 for extra time in schools.

1 "Districts" here refers to the the provincial education department structure at district level.

3 Theoretical framing

3.1 Formative assessment

In a discussion paper on "strengthening of education systems" (Gottelmann-Duret & Bahr, 2012), UNESCO reported that countries or schools that have made significant moves towards the "quality" sub-goal of the Sustainable Development Goals (SDGs) tend to focus on "performance, its assessment, and support" for both teachers and learners. In the majority of such countries or schools, there is an almost 50-50 balance between external standardised assessments and international assessments, and average performance shifts from "great to excellent" (34-35). Focusing on performance, assessment, and support necessarily requires effective use of assessment for learning. Teachers cannot respond to the assessment outcomes to adequately support learners with their specific difficulties, in the absence of formative assessment practices.

Although different users define formative assessment variously, the definition that has been adopted in this chapter is taken from the Council of Chief State School Officers (USA):

> *Formative assessment is a process used by teachers and students (learners) during instruction that provides feedback to adjust ongoing teaching and learning to help students improve their achievement of intended instructional outcomes*
>
> (CCSSO 2008, 3).

Key strategies of formative assessment include prompt feedback to both the teacher and the learners, learner-peer assessment, and self-assessment, all of which serve to improve teaching and enhance learning while the lesson is underway.

Formative assessment is often contrasted with summative assessment where the latter is defined as an assessment that is administered at the end of either a topic, a grade, or a project that learners have completed (Kibble 2017). The distinction between the two types of assessment is often blurred by how the assessment results are used. Moloi and Kanjee (2018) explain how to use summative assessment formatively.

Substantial published research on the efficacy of formative assessment and its impact on learning outcomes in terms of effect sizes that were as high as 0.7 of a standard deviation, is reported by Black and Wiliam (1998). Although the exact effect sizes from Black and Wiliam (1998) have been questioned on methodological grounds (Dunn & Mulvenon 2009), subsequent research has corroborated the findings that the use of formative assessment does enhance the achievement of learning outcomes. Studies on the strategies and the efficacy of formative assessment across school systems and contexts have been replicated by various researchers with similar findings. In South Africa, Kanjee (2020) reported significant improvements in learning outcomes, which he attributed to professional development that focused on training teachers on how to plan and implement formative assessments to enhance the chances of meaningful learning for all learners.

The JumpStart intervention included elements that are similar to other interventions such as the NumberSense Workbooks and teaching assistants but the use of tablets for the rapid real-time turnaround of assessment data for formative use in classes is a unique feature of this intervention.

4 Research design

The research questions answered in this chapter are:

1. How many addition and subtraction calculations can be answered correctly in two minutes in the JumpStart foundation phase classrooms?
2. What is inside the black box of the JumpStart intervention?

To answer the first question, we used the same EGMA data drawn on by the Roberts and Moloi (2021) study which was collected from 18 no-fee schools selected by the Ekurhuleni South district officials as poor performing. In each selected school, in each of Grade 1, 2, and 3, samples of thirty learners, or all the learners where there were less than thirty learners in a grade, were randomly selected (sequenced by height and selecting every fifth learner) to take the Core EGMA test under the guidance of a trained JumpStart teaching assistant. The JumpStart intervention was phased in to the 18 schools over three phases as shown in Table 1.

Table 1 Phased implementation of JumpStart intervention 2016-2019

	2016	2017	2018	2019
Phase 1: Schools 1-5	Baseline	Year 1	Year 2	Year 3
Phase 2: Schools 6-10		Baseline	Year 1	Year 2
Phase 3: Schools 11-18			Baseline	Year 1

Source: Data collected in the JumpStart intervention

The data from the baseline (outset of implementation) was collated for each grade. This was repeated for the EGMA data from after the first, second, and third years of implementation.

Table 2 Sample size for EGMA core items by Grade level

	Baseline (18 schools)	Year 1 (18 schools)	Year 2 (10 schools)	Year 3 (5 schools)
Grade 1	500	438	316	146
Grade 2	479	520	428	150
Grade 3	441	516	436	149
Total	1,420	1,474	1,180	445

Source: Adapted (to include number of schools) from Roberts & Moloi (2021, 133)

The cross-sectional quasi-experiment study, reported in Roberts and Moloi (2021), made small adjustments to the EGMA raw score based on the time of year of the EGMA administration, by making use of control school baseline and year 1 data. These shifts are reported in detail in Roberts and Moloi (2021).

In this chapter, we analyse the EGMA data in a different way. We draw on data at an item level for a testlet, and offer mathematical benchmarks drawn from the baseline (n = 1420) and year 3 (n = 445) [2] intervention of the JumpStart EGMA dataset. All of the EGMA assessments were conducted in the first two terms of the academic year, with most being administered in March or April. In this study, as a general benchmark was explored for the 'first term of a grade', fine-grained shifts relating to the exact month of EGMA administration were not made.

To do this we first needed to understand the EGMA test design. Platas, et al. (2014) offer detailed explanation of the core EGMA's development, descriptions of its technical adequacy, evidence of validity and reliability, as well as processes for local adaptation and training of assessors. The Core EGMA is a battery of six testlets, one of which is simple addition and subtraction.

The core EGMA was administered annually in Grades 1, 2, and 3, as noted in Roberts and Moloi (2021), in the language of learning and teaching for mathematics at Foundation Phase. The translations pertain to the instruction prompts and word problems. The language of learning and teaching in the JumpStart cohort was varied; the EGMA was administered to children in the language of teaching and learning at the school. In the urban environment of Gauteng this included: English, isiXhosa, isiZulu, Sepedi, Sesotho, Tshivenda, and Xitsonga.

Building on the work of Spaull and Ardington (in this volume) we chose to focus on simple addition and subtraction. Platas et al. explain that:

> *The Addition and Subtraction Level 1 subtests are timed tests (60 seconds) consisting of 20 items each that increase in difficulty. No addends are greater than 10, and no sums are greater than 19. The subtraction problems are the inverse of the addition problems. Three of the items mirror three of the Word Problems items. Assessors also keep track of whether the student used one of three problem-solving strategies: finger/tick marks, paper and pencil calculation, or mental arithmetic*
>
> (2014, 12-13).

The following exemplars are offered for simple addition and subtraction:

2 The 'baseline' refers to the first year in which JumpStart was working with a school. 'Year 3' refers to the third year of JumpStart intervention, when the programme was embedded in the school.

Figure 1 Exemplar EGMA core items for 'Simple addition and subtraction (level 1)'

Sample addition and subtraction level 1 items

1 + 3 = □	4 – 1 = □
3 + 2 = □	5 – 2 = □
6 + 2 = □	8 – 2 = □
4 + 5 = □	9 – 5 = □
3 + 3 = □	6 – 3 = □

Source: Reproduced on the basis of an original work published by RTI International and licensed under the Creative Commons Attribution 4.0 International License

Notice that all of the exemplars are single-digit addition or subtraction, result-unknown calculations with sums of no more than 10. The items stay in the same format (result-unknown calculations involving additive relations) but get progressively more difficult as children are required to bridge the ten, although the addends are smaller than 10, and the sums are smaller than 20. An example of a more difficult item in this testlet would be: 8 + 4 = □ or 12 – 4 = □. Basically the 'Simple Addition and Subtraction Level 1' testlet comprises of 20 whole-number addition calculations in the number range 0 to 20, and 20 whole-number subtraction calculations also in the number range 0 to 20.

The administration of this EGMA testlet is significant for its interpretation. This testlet is designed to assess procedural fluency in simple addition and subtraction calculations. It is also timed. The measure of attainment is therefore one of speed and accuracy: how many simple calculations can be correctly answered in one minute? There are 20 items – not because children in Grade 1 are expected to answer all 20, or to 'pass' at 50% (by getting 10 correct), but rather to ensure that the top-performing students can differentiate themselves from others by being able to respond to more calculations. We thought that this test provided a good measure for teachers, as they could easily set their own testlets to see how many simple addition, and how many simple subtraction calculations a child could complete in one minute.

The EGMA assessment was administered orally with a JumpStart teaching assistant reading out the instructions to the learners. The sampled learners were grouped and accommodated in one classroom and each was given a tablet on which they completed the assessment. Their work on the assessment was preceded by some sample questions, where the trained teaching assistant ensured that every child could use the tablet. The learners then made their choice of answer on the tablet. The teaching assistant then moved on to the next question, read this out, and all learners answered. The EGMA script or guide was provided in English and the teaching assistant translated it, on the spot into the language of teaching and learning of the school. The teaching assistants were matched by their home-language to the language of learning and teaching of the school. The quality and extent of on-the-spot translation was however not monitored. This was noted as design and data-collection weakness as it meant that the tests were not standardised and tightly scripted for each language. However, as the EGMA was administered by the same test administrators, and in the same way at each data-collection point, the comparison in attainment across the schools is considered adequate.

First, we considered the baseline data (*n* = 1,420) from the JumpStart schools. This was EGMA data collected in the year in which JumpStart was introduced to the school. As the EGMA was administered in the first or second term, we took this to be a measure of 'normal' or what may be expected in a typical Ekurhuleni South no-fee school, where there is no mathematics intervention. Next, we compared this to the EGMA data collected three years later, treatment year 3 data (*n* = 445). Again the data was collected in the first two terms of the academic year. We took this data to represent 'what was possible', or an aspirational target to which schools could aspire.

We focused on the EGMA item-level data from the 'Simple Addition and Subtraction' testlet. For each learner we calculated how many addition calculations they answered correctly in one minute, and how many subtraction calculations they answered correctly in one minute. Adding the results, we obtained how many addition or subtraction calculations each learner answered in two minutes. For each group (Treatment baseline and Treatment Year 3) we calculated the mean and standard deviation for each Grade. In addition, to understand the distribution of the results within each group and grade we plotted a box and whisker graph, presenting a 5-point summary depicting the minimum, first quartile, median, third quartile, and maximum values. This gave a quick visual impression of the mean for each grade, but also of the distribution of values around the mean. As any Foundation Phase class is likely to have a fairly wide spread of attainment, we thought this offered a simple but clear benchmark for teachers.

To answer the second question we drew on qualitative data. First, the JumpStart project managers were interviewed and offered their description of the intervention, reflecting on what they thought made it work. Second, the third author was interviewed and described how they manage the teaching assistants and school and district level relationships, offering his insights into what he thought the main features of the JumpStart intervention were. Finally the JumpStart Numeracy tracker was reviewed and discussed by all three authors. The design of the NumberSense workbooks is discussed in some detail in Brombacher and Roberts (this volume).

5 Analysis and findings

The JumpStart intervention was implemented at the FP in township primary schools in the Ekurhuleni South education district. The focus of the intervention was on improving learner performance in early grade mathematics (Grades 1-3). The intervention schools were identified by the education district authorities and were largely schools that had experienced consistent underperformance in both provincial and national assessments. A staggered approach to the intervention was adopted whereby five schools were included in the intervention in 2016, another five in 2017, and finally eight schools were included in 2018 resulting in eighteen schools participating in the intervention over a period of four years between 2016 and 2020 (Roberts & Moloi 2021). Ekurhuleni South is a district in a mix of urban and semi-urban settings characterised by townships and informal settlements. The population density is high, unemployment rates are high, and schools tend to be overcrowded (Matshipi et al. 2017; Manganyi 2011). The majority of schools in Ekurhuleni South belong in quintiles 1-3 and they are mainstream no-fee schools.

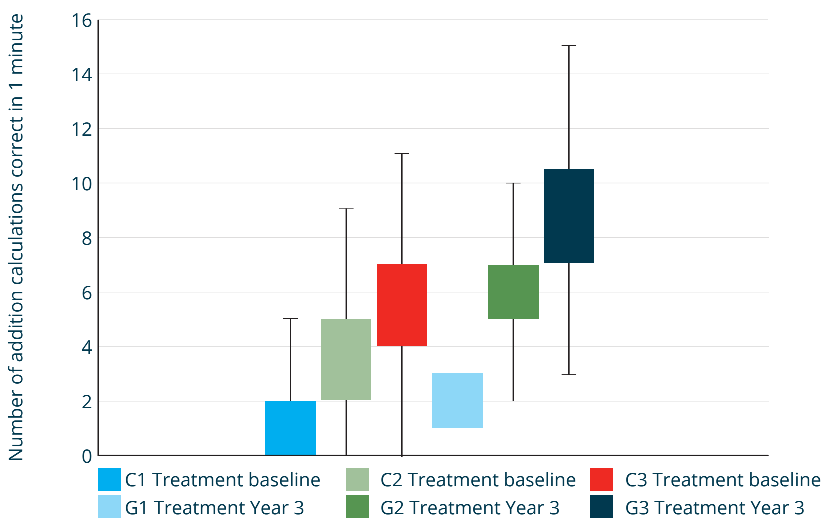
In this section we present our responses to the research questions:

1. How many addition and subtraction calculations can be correctly answered in two minutes in JumpStart Foundation Phase classrooms?
2. What is inside the black box of the JumpStart intervention?

5.1 EGMA benchmarks

Focusing only on the simple addition and subtraction EGMA testlet, we found that at the baseline, the average learner in the JumpStart schools was able to correctly answer one, four, and five addition sums in Grade 1, 2, and 3 respectively. Three years later when the JumpStart intervention was embedded into those same schools, the average learner could calculate two, six, and nine sums in Grade 1, 2 and 3 respectively.

Figure 2 **Addition calculations correct in 1 minute (by grade and group)**

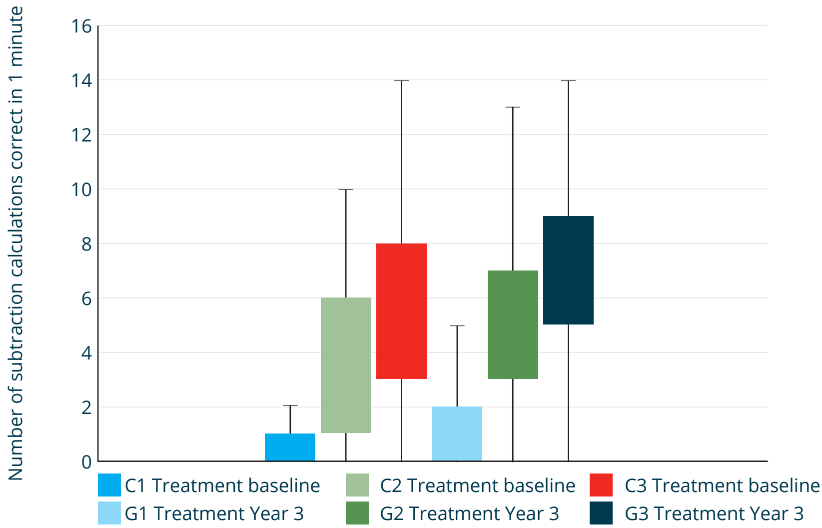


Source: Author created

Note that the 'G1 Treatment baseline' had a lower quartile and minimum of 0. One quarter of the children obtained zero correct addition calculations in one minute. For 'G2 Treatment baseline' and 'G1 Treatment Year 3', the minimum score was zero.

Using the baseline data the norm for subtraction calculations was again one, four, and five (Grade 1, 2, and 3 respectively). By year 3 this had increased to one, five, and seven (for the average Grade 1, 2, and 3 learners).

Figure 3 **Subtraction calculations correct in 1 minute (by grade and group)**

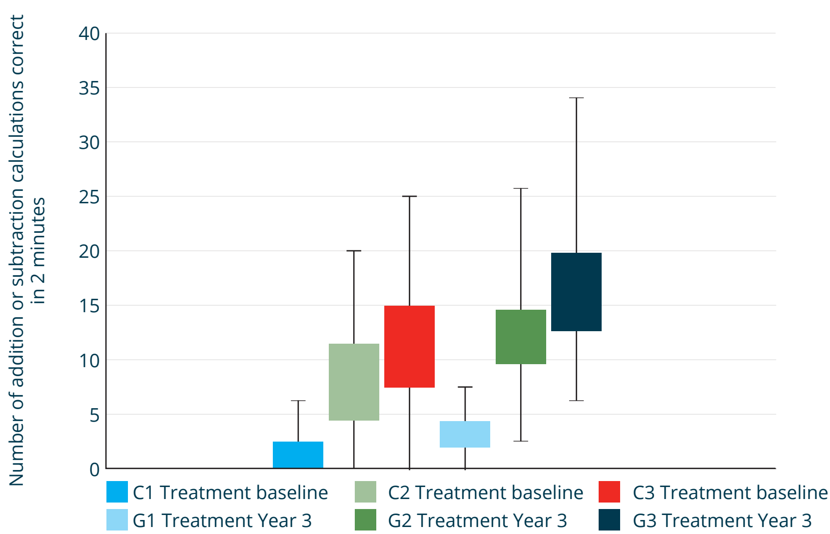


Source: Author created

Note that the all of the data sets had minimum scores of 0. In the case of 'G1 Treatment baseline', half of the children were not able to correctly answer any subtraction calculations in one minute. For the 'Grade 1 Treatment Year 3' group, one quarter of the learners could not answer any subtraction calculations in one minute.

Combining the scores for simple addition and simple subtraction results in the following:

Figure 4 **Addition or subtraction calculations correct in 2 minutes (by grade and group)**



Source: Author created

Once again the prevalence of children obtaining 0 (not being able to answer any addition or subtraction calculations) is notable. It is significant that the Grade 1 average score is low for both addition and subtraction. Given that the EGMA was administered in the first or second term, this is to be expected as Grade 1s are just entering formal school and not yet reading and writing numerals fluently. The changes in attainment are notable at Grade 2 and 3 levels where the learners in these JumpStart school were able to correctly calculate more than ten calculations in one minute (in the first half of Grade 2) and more than 15 in two minutes (in the first term of Grade 3). Besides providing a sense of what is likely (considering the baseline as normative), and what is possible with a focus on mathematics (considering the improvement after three years of JumpStart), it is also helpful to notice the distribution of results. At the baseline there were some Grade 1 children performing at the expected 'Grade 2 level' and this was evident in both groups. At Grade 3 level, there were children who were able to correctly complete more than 25 calculations in two minutes.

5.2 Inside the black box of the JumpStart project

The JumpStart intervention is conducted at district level.

5.2.1 District level coordination

JumpStart appoints a project coordinator who is assigned to be a liaison person between the project schools and the education district. An important requirement for being eligible as a coordinator is that the individual must be a teacher with some minimum teaching experience. The coordinator is responsible for ensuring that district officials such as subject advisors are kept informed about the activities and progress of the project in the schools. The project coordinator's responsibilities include monitoring teachers' attendance at scheduled workshops, keeping records of teaching assistants' attendance in the assigned schools and making sure that the schools receive NumberSense workbooks in time. Once appointed the coordinator introduces themselves to the project schools as a partner and co-workers. They must assure schools upfront about their independence from the formal echelons of the education system such as the district office and circuits. In this way, they gain the trust and confidence of both the school management teams (SMT) and the teachers, thus creating a supportive working environment.

The project coordinator ensures that the district gains access to the EGMA results from all schools that are in the intervention. This helps districts monitor school performance and identify schools that require support. The project coordinator and the district officials, who are usually subject advisors, have access to and share information on the district dashboard. Regular interactions help entrench the approach of the intervention into the reporting systems of the district and are thus encouraged for sustainability.

5.2.2 The intervention package

The JumpStart Project Manager described the JumpStart theory of change as comprising of five pillars: (a) teachers, (b) mentoring of teaching assistants, (c)

NumberSense workbooks, (d) pedagogy and (e) feedback via monitoring and evaluation (and using a mobile application). Looking across the five pillars, the project manager considered 'real-time monitoring and evaluation' facilitated via the mobile application as 'a key component', noting the importance of identifying areas of difficulty immediately through the formative assessment process. The project coordinator identified the NumberSense workbooks and the teaching assistants to be the key components of the JumpStart intervention.

A. Teachers and teacher training

Teacher training is carried out once a year. The content to be treated in the training is drawn mainly from the areas that school dashboards reflect as challenging to learners but may be augmented with other priority areas identified by either the schools or the districts. The focus of the training is on the pedagogy of teaching mathematics in the early grades (Grade 1-3) along the theoretical trajectories identified by Clements and Sarama (2014).

A teacher can view which books are being used by her class. Formative assessment and prompt feedback enable teachers to use the results to provide appropriate support at the level where individual learners are ready to learn new material. By reducing the workload on teachers, the use of teaching assistants in the JumpStart intervention creates desired space for teachers to focus on specialised professional responsibilities such as preparing and delivering pedagogically sound lessons.

B. Mentoring and teaching assistants

Teaching assistants see each Foundation Phase maths class twice per week. From the perspective of the teacher and learner, two 40-minute maths lessons per week are NumberSense lessons. The teaching assistant supports and monitors learners whilst they complete their tasks in their workbooks.

Teaching assistants are appointed and deployed in each project school so that there is at least one teaching assistant for each grade. Appointees are unemployed young adults, a socio-economic phenomenon that appears to be on the increase in South Africa (Oluwajodu et al. 2015). The teaching assistants are given task-specific training by a senior professional employed by JumpStart and are also given regular support on-site. They are given a stipend which alleviates the impact of unemployment to some extent.

The use of teaching assistants is one major mechanism for change and enhancing efficiency in schools. One of the systemic challenges in classrooms is that teachers often get overwhelmed by administrative and clerical work and this tends to compromise their focus on actual teaching. The use of teaching assistants turned out to be strategic in this regard. The responsibility of teaching assistants is mainly twofold. Firstly, they guide and help learners through the NumberSense workbooks. Secondly, they collect important assessment data for real-time feedback to both teachers and learners for formative pedagogy.

The main function of teaching assistants is to take on most of the administrative responsibilities of teachers such as marking learners' workbooks, supervising learners as they work through the workbooks, administering EGMA tests with learners, providing support to individual learners, and guiding learners to relevant workbooks and sections of workbooks that they need to focus on, based on identified strengths and weaknesses obtained from the assessment results.

C. NumberSense workbooks

JumpStart provides every learner with three graded NumberSense workbooks. The workbooks are underpinned by a strong theoretical basis and structured to reflect how children learn mathematical concepts, which has been described in some detail in Brombacher (2019) and Brombacher and Roberts (this volume).

Figure 5 **Tracking learner progress by workbook**

Source: Screen shot from JumpStart application

Learners work regularly and sequentially from page to page in a workbook and from one workbook to the next, depending on whether there is evidence of mastery of work in all the sections and concepts in the previous workbook. A learner who has not mastered all the sections and concepts of a workbook receives individual remedial support in the affected sections of the workbook while others who have shown mastery are allowed to proceed to the next workbook. This individual attention, which is almost impossible for teachers without the help of teaching assistants, affords each learner the opportunity to learn at their own pace and develop concepts along the mathematics trajectories.

D. Pedagogy

The underlying pedagogy of the NumberSense approach to mathematics is described in some detail by the programme designer in Brombacher and Roberts (this volume). Of interest in this chapter is how this pedagogy is interpreted by the JumpStart implementers. The JumpStart project manager explained his understanding of the mathematics pedagogy of the NumberSense workbooks:

> *Mathematics is a problem-solving activity, and in the early grades, we want to teach learners to count with meaning not by rote. We want them to manipulate numbers, and we want them to start solving the problems…*

(Roberts 2019, 142)

The project coordinator explained that the NumberSense workbooks do not cover all the curriculum content but they cover key competencies in NumberSense (Number, Operations and Relationships) which constitutes the largest component (more than 60%) of the Foundation Phase mathematics in the South African Curriculum and Assessment Statement (CAPS). He further explained that the workbooks are graded by level of complexity of the skills that learners are expected to demonstrate in the NumberSense tasks. The NumberSense tasks support the CAPS content areas – including "Number, Operations and Relationships", "Patterns Function and Algebra" and "Measurement". Teachers continue to teach three mathematics lessons per week for CAPS lessons and coverage, and students consolidate this work – at the right level – through the two NumberSense lessons per week. As such learners are still exposed to the same number of mathematics lessons (five per week) as in other schools. Their mathematics lessons are however split into three teacher-led CAPS lessons for new content and ideas, and two teacher assistant-led Number Sense consolidation lessons.

A key part of the pedagogical approach is the real-time monitoring and rapid feedback on learner progress and their areas of difficulties. The project manager explained:

> *We do not want to wait for the end of exams, we want to do it immediately. We want to gather [feedback] this week so we can start making an impact the next week.*
>
> (Roberts 2019, 142)

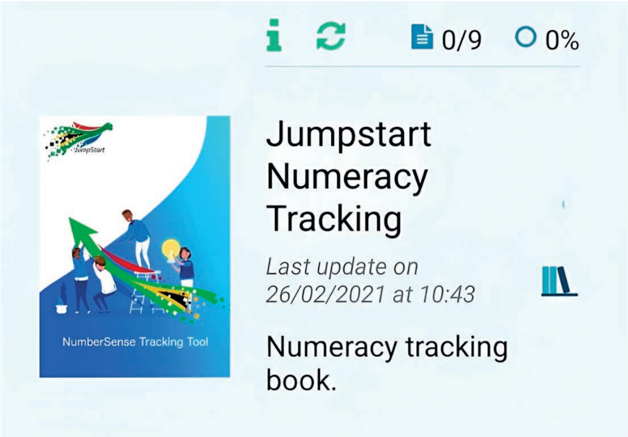
This explicit focus on using formative assessment to inform learning is discussed below.

E. Feedback via monitoring and evaluation (formative assessment)

The JumpStart team refers to one of its key pillars as "feedback via monitoring and evaluation". In our view, JumpStart provides an example where summative assessment is used formatively, as advocated by Moloi and Kanjee (2018). Standardised EGMA assessments are administered once per year for a randomly selected sample of learners. These results are analysed and presented to district officials, school managers, and teachers so that they can focus their subsequent teaching to areas of concern. In addition, the use of tablets makes the results available in real time so when a learner completes a page in a workbook, these formative results are immediately fed back to both the teacher and the learners for immediate formative use and remediation.

To embed formative assessment into the teaching process and enhance data-driven support to learners, JumpStart provides each teacher assistant with a tablet which includes both the EGMA assessment instrument and the full suite of NumberSense workbooks on a page-by-page basis and a database of learners' particulars.

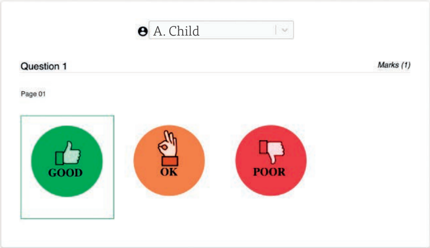
Figure 6 **JumpStart Numeracy tracking book (online app used by teaching assistants)**



Source: Screenshot of JumpStart Numeracy tracking application

Teaching assistants use the JumpTrak platform (a tablet-based application) to track learner progress through the NumberSense printed workbook. In each NumberSense lesson, the teacher assistant records which page in which NumberSense book the learner is working (this provides a page-completion rate per week and per term).

Figure 7 **Teaching assistant rating a child on a particular page**

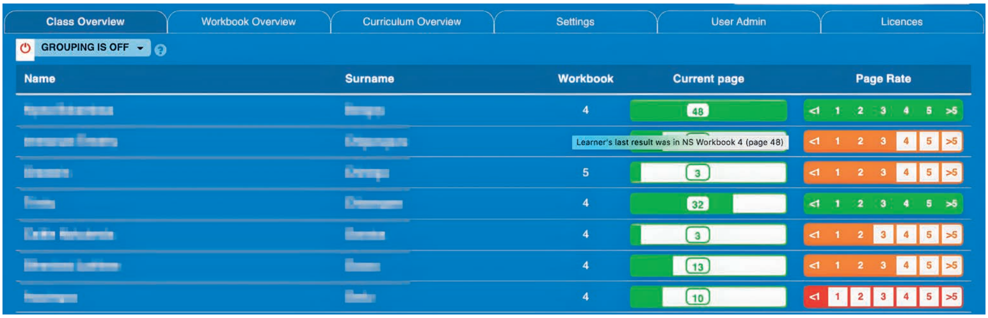


Source: Screenshot from JumpStart Numeracy tracking application

In addition, the teacher assistant makes a qualitative judgement on how each learner is managing with the mathematics content on that page: Green denotes 'good', which means the learner is getting most of their tasks on the page correct; yellow denotes 'OK', which means there are some errors, and red denotes 'poor', which means the learner's work on this question has many errors. In the next NumberSense lesson the learners who are coded red are called together for individual remediation on the work on that page.

The teacher and teaching assistant also have a whole-class view, where they can see the page-by-page progress of their entire class:

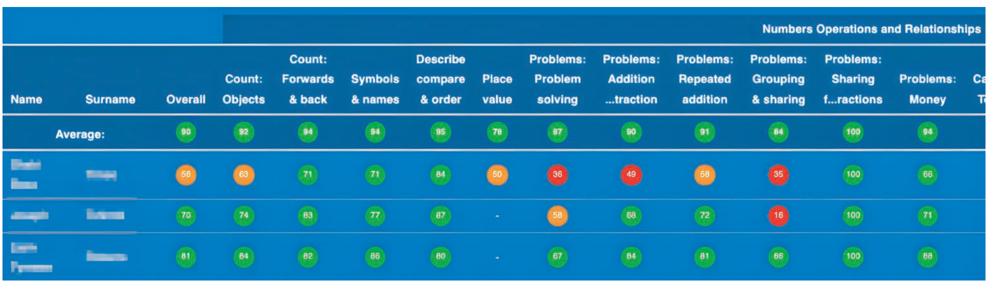
Figure 8 **JumpStart Teacher interface to see the progress of a class (workbook and page view)**



Source: Screenshot of JumpStart Numeracy tracking application

By using this view teaching assistants can quickly identify the children requiring specific support. By switching to the curriculum view, they can see the children in their class, and identify which children require support on particular topics.

Figure 9 **JumpStart Teacher interface to see the progress of a class (curriculum view)**



Source: Screenshot of JumpStart Numeracy tracking application

In the Foundation Phase class, the children who have red dots 'grouping and sharing problems' can be called to the teacher's desk or onto the mat, to work on that concept as a small group.

The teacher assistant enters individual learner responses on the tablet while the lesson is in progress. The engagement of teaching assistants, combined with their use of tablets to collect and upload learner-assessment data in real-time, significantly reduces turnaround times for teachers to give prompt feedback to individual learners while the assessment evidence can still be acted upon. The Numeracy tracker eases the administration of formative assessment in the mathematics lessons as the children, topics, and pages or questions requiring more attention are quickly collated and clearly visible. A learner gets feedback and support at the time of need when they are best ready to learn, a pedagogy promoted in Black and Wiliam (2009).

6 Conclusion

This chapter extends the evidence of promising improvement in EGMA attainment reported by Roberts (2019) and Roberts and Moloi (2021) for learners in 18 JumpStart schools in Ekurhuleni South. It provides evidence of the specific learning gains which have been possible in JumpStart schools – focusing only on the EGMA Addition and Subtraction level 1 testlet.

At the baseline, the average Grade 3 learner in a JumpStart school was only able to answer ten addition and subtraction calculations correctly in two minutes. After three years of JumpStart intervention, the average Grade 3 learner was able to correctly answer 16 simple addition and subtraction calculations. This provides a preliminary EGMA benchmark of what may be 'normal' (at baseline, prior to an intervention) and what may be possible / aspirational.

Table 3 Number of correct calculations in timed EGMA Addition and Subtraction level 1 testlet

	Likely / normal (JumpStart baseline)			Aspirational/ possible (Jumpstart after 3 years)		
Addition and subtraction level 1	**Addition (1 min)**	**Sub-traction (1 min)**	**Addition and sub-traction (2 mins)**	**Addition (1 min)**	**Sub-traction (1 min)**	**Addition and sub-traction (2 mins)**
Grade 1	1	1	2	1	2	3
Grade 2	4	4	8	5	6	11
Grade 3	5	6	11	7	9	16

Source: Author created

Increased attainment in the EGMA as whole is now made more meaningful to teachers in terms of calculation fluency in the timed Addition and Subtraction level 1 testlet.

Before considering taking the JumpStart intervention to a wider scale, it is worth trying to open the 'black box' of the JumpStart intervention to describe the key inputs and processes, which are thought to improve teaching and enhance the learning of early grade mathematics in JumpStart schools. We think that there are particular project inputs which are critical to the JumpStart intervention:

1. District level coordination;
2. Teaching assistants who are timetabled to teach two of the normal mathematics lessons per week for each class in a grade (Grades 1 to 3) and are mentored;
3. NumberSense workbooks designed for individual learner consolidation at the right level of their mathematical development
4. Formative assessment making use of the JumpStart Numeracy tracker, which is used by a teaching assistant to administer and respond to quality formative assessment processes.

While 'pedagogy' and 'teachers and teacher training' are also considered to be pillars of the JumpStart offering, these are not described in detail by the JumpStart project managers and coordinator. Their combination of carefully designed workbooks – which are given to the children for teaching at the right level of mathematics – coupled with the JumpStart Numeracy tracker (used by a teaching assistant with a tablet) provides a powerful yet simple, mechanism for differentiated pacing in a mathematics classroom. In a 'one-size-fits-all' model the pace of the whole class is tracked against curriculum expectations, the curriculum – and what is being taught – is in focus and is policed by teachers, heads of department and subject advisors. With the JumpStart Numeracy tracker the focus shifts from teaching to learning. What is tracked is not curriculum coverage, but rather the individual child's learning. It is this shift in attention on what is tracked that allows for differentiated pacing.

The JumpStart intervention is showing promise on a small scale of 18 schools at a district level. Its intervention design and key ingredients are well documented. It therefore seems appropriate that the promising findings evident on a small scale now be tested more rigorously with a randomised control trial on a larger scale.

Acknowledgments

The JumpStart intervention is a partnership between the Gauteng Department of Education, the Ekurhuleni South district, and JumpStart and is funded by the Michael & Susan Dell Foundation.

References

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5–31.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: principles, policy & practice*, 5(1), 7–74.

Brombacher and Associates (2019). *Making the most of NumberSense workbooks Available at:* https://www.numbersense.co.za/assets/images/file_download/Workshop_Making-the-most-of-the-NSWKBKs.pdf, (Accessed 30 March 2022).

Clements, D. H., & Sarama, J. (2014). *Learning and teaching early math: The learning trajectories approach*. London: Routledge.

Council of Chief State School Officers (2008). Attributes of Effective Formative Assessment: A work product coordinated by Sarah McManus, North Carolina Department of Public Instruction for the CCSSO Formative Assessment for Students and Teachers (FAST) SCASS. Washington, D. C.: CCSSO.

Desimone, L. M., & Pak, K. (2017). Instructional coaching as high-quality professional development. *Theory into practice*, 56(1), 3–12.

Dunn, K. E. & Mulvenon, S. W. (2009). A critical review of research on formative assessment: The limited scientific evidence of the impact of formative assessment in education. *Practical Assessment, Research and Evaluation*, 14(1), 7.

Kanjee, A. (2020). Exploring primary school teachers' use of formative assessment across fee and no-fee schools. *South African Journal of Childhood Education*, 10(1), 1–13.

Khosa, G. (2014). Lessons on District-level Support and Integration. In Khosa, G. (Ed.) *Systemic School Improvement Interventions in South Africa: Some Practical Lessons from Development Practitioners*. Johannesburg: JET Education Services.

Kibble, J. D. (2017). Best practices in summative assessment. *Advances in physiology education*, 41(1), 110–119.

Manganyi, J.N. (2011). How do public high schools in Ekurhuleni South District (Gauteng) implement educational reform? A mini-dissertation submitted in partial fulfillment for a Master of Public Administration degree. University of Limpopo.

Matshipi, M.G., Mulaudzi, N.O., & Mashau, T.S. (2017). Causes of overcrowded classes in rural primary schools. *Journal of Social Sciences*, *51*(1–3), 109–114.

Moloi, M., & Kanjee, A. (2018). Beyond test scores: A framework for reporting mathematics assessment results to enhance teaching and learning. *Pythagoras*, 39 (1), 1–14

Mouton, J., Wildschut, L., Richter, T., & Pocock, R. (2013). Review project: Final report (Unpublished report). Johannesburg, RSA: Zenex Foundation.

Oluwajodu, F., Greyling, L., Blaauw, D., & Kleynhans, E. P. (2015). Graduate unemployment in South Africa: Perspectives from the banking sector. *SA Journal of Human Resource Management*, 13(1), 1–9.

Platas, L. M., Ketterlin-Gellar, L., Brombacher, A., & Sitabkhan, Y. (2014). *Early Grade Mathematics Assessment (EGMA) toolkit*. Research Triangle Park, NC: RTI International.

Roberts, N. (2019). Assessing Early Grade Mathematics Learner Outcomes Using m-Learning. In *Interactive Mobile Communication, Technologies and Learning* Springer: Cham.

Roberts, N. & Moloi, Q. (2021). JumpStart Programme Impact on the Mathematics Learning Outcomes: A cross sectional quasi-experiment using the Early Grde Mathematics Assessment In Qhobela, M., Ntsohi, M.M.E., & Mohafa, L.G. Eds. *Book of Proceedings of the 29th Annual Conference of the Southern African Association for Research in Mathematics, Science and Technology Education*. Lesotho: SAARMSTE.

SACMEQ (2017). The SACMEQ IV Project in South Africa: A Study of the Conditions of Schooling and the Quality of Education. Available at: http://www.sacmeq.org/sites/default/files/sacmeq/reports/sacmeq-iv/national-reports/sacmeq_iv_project_in_south_africa_report.pdf (Accessed 12 July 2020).

Van der Berg, S. (2008). How effective are poor schools? Poverty and educational outcomes in South Africa. *Studies in Educational Evaluation*, 34(3), 145–154.

16

What works and what scales? Returning to a tradition of evidence-based system-wide programmes

STEPHEN TAYLOR & NIC SPAULL

KEYWORDS
system-wide reform, interventions, scale-up, early grade reading and mathematics

Abstract

International cases, and South Africa's own experiences of system-wide education reforms, demonstrate that it is possible to make significant progress in reading and mathematics outcomes at a national scale. This chapter draws together the lessons gleaned from early grade reading and mathematics interventions implemented between 2010 and 2022 and argues that a new round of sustained, system-wide interventions focused directly on classroom teaching and learning is urgently needed. If paired with a strengthened accountability for learning, also at a systemic level, these interventions can provide teachers and learners with the tools and opportunities needed to reduce the inequalities in foundational learning.

1 Introduction

The compilation of chapters in this book (and indeed across all three volumes) signifies several important developments within the South African education sector. A range of actors within the sector are now working in closer collaboration than before, with partnerships increasingly spanning different disciplines and sectors. The authors in this book include education academics, linguists, economists, materials developers, and practitioners, and work across academic institutions, non-governmental organisations (NGOs) and government. This collaboration is reflected in the way the interventions are covered: detailed descriptions of the pedagogical principles embedded in the intervention design and materials are presented along with attempts to measure the impacts of interventions using statistical methods.

Another positive shift is that many of the people administering educational interventions recognise the potential to inform wider policy and practice rather than only impacting the direct recipients of their interventions. As a result, we now know a lot more about 'what works' and 'what scales', and this chapter aims to consolidate these lessons. We argue that education policymakers should build on these lessons and return to a South African tradition of implementing system-wide reforms, rather than only in localised 'pockets' which has become the norm. Where the decade began with bold and research-based reforms to the curriculum (CAPS), the introduction of a universal assessment system (ANAs) and a new national workbook programme, the norm is now for 'new interventions' to be implemented in a few schools only or by parastatal providers in a few thousand schools. While experimentation is necessary, and our argument is not that any new idea should simply be 'taken to full scale', we argue that evidence-based system-wide implementation of new pro-poor reforms (rather than a proliferation of projects, pilots, partnerships, and priorities) is ultimately how South Africa can return to its pre-pandemic trajectory of improvement.

2 The role of 'interventions' in country-level improvements

Before reflecting on the specific interventions covered in this volume, we might first question whether the modality of 'interventions' is a promising approach to addressing the educational challenges of a country like South Africa. Is it wise to try to improve education by scaling up a few promising interventions, each focusing on one particular outcome (e.g. reading), or should one rather work on the system as a whole? The latter might place greater emphasis on coordination between departmental silos, strengthening accountability, or reducing learner-educator ratios. One possible concern with a system-strengthening approach is that the impact on specific learning outcomes may be frustratingly slow and indirect. Furthermore, many system-strengthening approaches tackle less controversial issues such as free school meals or the provision of materials, rather than changing pedagogical practices or teacher evaluations. This is one of the core arguments made by a number of authors in a new edited volume *Schooling for All* (Sandefur 2022) who distinguish between 'what works' (improves learning) and 'what scales' (easy to implement).

Although one might be impatient with a generic system-strengthening approach, the reality is that interventions also typically have frustratingly small effects relative to the magnitude of existing inequalities in learning outcomes. For example, the gap between no-fee schools and fee-paying schools in the South African TIMSS (Grade 5) assessment of 2019 was equivalent to approximately three years' worth of learning (105 points)[1], whereas the impact of the most effective EGRS coaching intervention was equivalent to about one seventh of that difference (an extra 40% of a year of learning). The limited impact of interventions is perhaps inevitable given what we know about the overall determinants of education outcomes. For example, research consistently demonstrates that a child's socio-economic background is the largest factor in determining learning outcomes (e.g. Taylor and Yu 2009). Most interventions would aim to improve the quality of teaching, but this would be to make a marginal impact on only one factor that influences a child's learning.

Another way to approach the issue is to consider what has caused education systems to improve around the world. Bruns et al. (2012) describe education system improvement in Brazil and point to the role of reforms to equalise funding and protect educational opportunities for poorer children (already major policy imperatives in South Africa), a system of national testing to make school performance information widely known, and the continuity of these reforms across different government administrations. London (2021), describing education improvement in Vietnam, also highlights the importance of sustained political commitment and of multiple mutual layers of accountability. Interestingly, this phrasing is similar to what the National Development Plan recommends for improving education: a system of "results-oriented mutual accountability" (NPC 2012).

Recent research from other middle-income countries reveals both that learning improvements are possible at the national level, but also that the political economy of learning and quality reforms to reach those improvements was always contested, turbulent, and non-linear, as Bruns et al. (forthcoming) argue in their paper "Technocrats and Turbulence":

> *Peru is part of an Andean wave of system-wide reforms aimed at raising teacher quality, with major reforms in Colombia 2002, Ecuador in 2009, and Chile beginning in 2004 and culminating in a comprehensive reform in 2016. All these systems have seen improvements in learning…Peru stands out for having achieved the largest learning gains through 2020.*

Yet even in Peru the improvement journey was not straightforward:

> *Education policies have also been markedly affected by the country's general institutional weakness and deep patterns of corruption, which have translated into constant ministerial changes and discontinuities in policy agendas that have either hampered or slowed the pace of reforms*
>
> (Balarin 2021, 4).

1 In 2002, nationally representative samples of both Grade 8 and Grade 9 learners wrote the same TIMSS assessments, creating a rare opportunity to measure a grade-level of learning in terms of the TIMSS scale. The difference between South Africa's Grade 8 and 9 average scores was 21 points for mathematics and 22 points for science. Meanwhile, an HSRC (2012) report on TIMSS results suggests that a grade-level difference is about 40 points on the TIMSS scale. We assume that the true value probably lies somewhere in between 21 points and 40 points, meaning that a gap of 105 points is roughly equivalent to about three years of learning.

In most improving countries, teaching and learning improve over time through a process of 'two steps forward one step back' as politics and priorities change within the country. Examples of an unbroken chain of consecutive incremental improvements are the exception (e.g. Ceará in Brazil) rather than the rule. Nevertheless, that some institutional similarities between South Africa and Peru exist, and that improvement is possible amidst contestation and disruption, provides hope that such national reforms are indeed possible in the future in South Africa.

While educational improvement is inherently political, technical aspects like pedagogy and resources are important. Drawing on the successes of Sobral and Ceará in Brazil, Tusome in Kenya, and "Teaching at the Right Level" in Zambia, Evans (2022) concludes that structured pedagogy support and providing teachers and learners with the right learning-support materials are key ingredients to deliver systemic impact.

Putting all this together, while fundamental system reforms are clearly the driving force behind national improvements in outcomes, it seems necessary also to have the tools and support provided by the sorts of interventions discussed in this book. There is also likely to be a positive interaction between capacity-based interventions and accountability measures (Spaull 2015). As the chapter by Mckay and Spaull implies, the DBE Workbooks intervention (the only full-scale national intervention in this book) attributes its success in part to the simultaneous curriculum reforms and introduction of the Annual National Assessments. These three reforms have also been identified by the Department of Basic Education (2020) as the most likely explanation for South Africa's improvements in international testing programmes seen from 2011 onwards. Even an intervention providing materials only (DBE Workbooks), when sustained over a long period of time and aligned to a system of measuring learning outcomes (soft accountability), appears to have significantly changed instructional practice across the system and contributed to improved learning.

This brief review of what causes country-level improvements in learning shows that the debate between interventions and system-strengthening is ultimately a false dichotomy. If a multitude of interventions are implemented in parts of the country, temporarily and without alignment to broader curriculum policy, they will not achieve significant system-wide gains. However, if implemented across the system for a sustained period in coherence with accountability for learning outcomes, they can be a significant factor contributing to improvements over time.

If we accept that interventions will only be effective if implemented in this systemic fashion, it also becomes clear that the appropriate purpose for any interventions that are not implemented at full scale is to generate evidence to inform system-wide support. Even interventions that are implemented across several thousands of schools will not achieve lasting gains at the national level and should therefore be designed with the explicit purpose of generating relevant learning. We need fewer interventions that are neither at scale, nor generating evidence to inform the system.

Based on an analysis of historical trends in country-level improvements in international assessments of learning, Gustafsson (2019) argues that the fastest possible country-level improvement in the percentage of children who reach acceptable benchmarks of learning proficiency is two percentage points a year. This implies that if South Africa were to improve reading proficiency at this so-called 'speed limit', it would take at least 30 years to increase the percentage of Grade 4 children reading with acceptable comprehension in the PIRLS assessment from 22% to 80%.

How does this aspirational level of country improvement compare to the potential impacts of some of the interventions discussed in this book? The most consistent and positive causal effects were observed in the EGRS and Funda Wande interventions, which were structured learning programmes consisting of lesson plans, integrated reading materials (including decodable texts) and support for teachers, either through coaching or through teaching assistants. The EGRS coaching intervention increased the percentage of Grade 2 children reaching the Grade 2 Setswana benchmark of 40 words-correct-per-minute from 33% (in the control group) to 43% (in the coaching group) (Spaull & Taylor, current volume). This improvement of ten percentage points in two years is faster than the 'speed limit' of national improvement, but this is understandable since the intervention was administered in only 50 schools. If this type of programme was implemented on a national scale, there would likely be some loss of impact. On the other hand, if aligned to a national focus on reaching critical reading benchmarks in home-language reading, this sort of intervention could play an important role in driving national improvements at the speed limit.

3 A decade of South African early grade interventions: Lessons learned

South Africa today is in the fortunate position of having a large body of empirical research on early grade reading and mathematics interventions that did not exist ten years ago. Many of the most important of these interventions are documented in this book, facilitating a bird's eye view of their similarities and differences. Table 1 categorises the interventions in this book according to six criteria: whether they targeted reading or mathematics, the implementation modalities used, the level of scale, whether there was a plan at the outset of the intervention to measure impact on learning, the evaluation methods that were used, and the size of the effect (if valid causal inference was possible).

3.1 Lessons about why interventions work

A similar number of reading and mathematics interventions are covered in this volume, but there is more consensus around how to improve reading than mathematics. This is because we have more studies that measure causal impacts on reading outcomes than mathematics outcomes, and because the findings emerging from these reading interventions have converged in support of one general approach: structured learning programmes that include lesson plans, integrated reading materials including decodable graded readers, and in-classroom support for teachers. Putting together what worked and what did not work across the various EGRS interventions (Fleisch & Alsofrom, this volume), and considering the Funda Wande evaluations, it is clear that these elements of a structured learning programme have now been proven to work across a range of South African contexts.

Table 1 Characterisation of interventions in this book

	Reading or Maths	Type of intervention	Level of scale	Designed to measure impact?	Evaluation methods	Causal effect size
The development and implementation of the DBE Workbook Series (literacy and mathematics)	Both	Material only	Full scale	No	Materials reviews	N/A
A Decade of the Early Grade Reading Study (EGRS): Lessons Learned and an Agenda for Action	Reading	Structured Pedagogy: Materials and Coaching	Small	Yes	RCT, lesson observations, case studies	Moderate
Early Grade Reading and the National Education Collaboration Trust (NECT): 2013-2020	Reading	Materials plus cascade training	Large	No	Descriptive statistics	N/A
The Science and Magic of reading in Action: A decade of Room to Read in South Africa (2010-2020)	Reading	Structured Pedagogy: Materials and Coaching	Small	No	Descriptive statistics	N/A
Vula Bula: Conceptualisation, Development and Implementation & Evaluating the impact of Vula Bula Anthologies in the Eastern Cape	Reading	Material only	Large	No	Quasi-experimental	Moderate
Teacher Assistants and Early Grade Reading in Grade 1: The case of the Funda Wande Sepedi Intervention in Limpopo 2021-2022	Both	Structured Pedagogy: Materials and TA's	Small	Yes	RCT, lesson observations, case studies	Moderate
Using technology to improve English First Additional Language (EFAL): The case of 'Reading Eggs' in South Africa 2012-2020	Reading	Ed tech – direct learner engagement	Small	No	Descriptive statistics	NA (large gains described)

	Reading or Maths	Type of intervention	Level of scale	Designed to measure impact?	Evaluation methods	Causal effect size
The conceptualisation, development, implementation and evaluation of the Grade R Mathematics Project (R-Maths) in the Western Cape (2016-2019)	Maths	Materials plus Training	Medium	No	Quasi-experimental	Small
The implementation of the Bala Wande programme in Grade 1 in three provinces: lessons learned	Maths	Structured Pedagogy: Materials and Coaching	Small	Yes	RCT, lesson observations, case studies	Moderate
Improving mental mathematics in the Foundation Phase	Maths	3-week programme (8 lessons) using online training	Medium	No	Descriptive statistics	N/A (large gains described)
Lessons Learned and Evidence of Impact: Formative Assessment in an Integrated Reading and Mathematics Intervention	Maths	Training workshops	Small	Yes	RCT, lesson observations	N/A (only changes in teacher practice reported)
NumberSense: A focus of Shikaya	Maths	Materials and coaching	Small	No	Descriptive statistics	N/A (large gains described)
The JumpStart intervention: Using Number Sense workbooks and formative assessment to improve learning outcomes in early grade mathematics	Maths	Ed tech, materials plus TA's	Small	No	Descriptive statistics	N/A (large gains described)

Even Funda Wande's successful teacher assistant programme used assistants in support of a structured learning programme, a point emphasised by Makaluza & Mpeta (this volume). There is also more evidence on how to improve Home Language (HL) reading than English as First Additional Language (EFAL). Furthermore, the EGRS studies showed that a home-language coaching programme positively benefited EFAL, but an EFAL coaching programme did not benefit home-language reading. Although language policy is not the focus of any chapters in this volume, it is worth emphasising that in 2022 almost all South African policy analysts and researchers would agree that children should be taught to read first in their home language (for at least Grades R to 3) rather than a 'straight-for-English' approach. While that was already formal government policy at the turn of the decade (2010), shifting global consensus on this topic, and new South African evidence on the topic (Taylor & Von Fintel, 2016) have meant that there is now universal agreement on this point.

Some of the chapters provide detailed explanations of the pedagogical principles and approaches that informed the design of interventions. The lesson that emerges from these chapters is that details matter. This book should not leave simplistic impressions like 'graded readers work' or 'teaching assistants are effective'. As one example, Katz & Rees (this volume) describe the development of the Vula Bula series of graded readers. These stories were not translated or versioned from English, but were designed specifically for each African language, cognizant of their particular orthographic features, such as word length, consonant density, and whether the language is written conjunctively or disjunctively. There is a risk that providing other graded readers not developed with a similar approach will not have the same positive impact that the Vula Bula books appear to have (see Ardington & Spaull, this volume). Similarly, Makaluza & Mpeta's chapter is instructive when considering the 'necessary conditions' for an effective teacher assistant programme. These include a rigorous recruitment process (with assessments of literacy, numeracy, and soft skills), a structured set of learning materials, close monitoring and effective support. Moreover, it is important to note how teacher assistants were used. The Funda Wande intervention used them to support a structured learning programme, rather than to conduct afternoon activities or as a general resource for teachers to use at their discretion. Similarly, the chapters describing the NumberSense workbooks (Brombacher et al.), the Mental Starters Assessment Project (Venkat & Graven) and the Learner Activity Books used in Bala Wande's interventions (Sapire et al.) all emphasise the importance of specific pedagogical design features.

3.2 Lessons about scale

Only one intervention covered in this book was implemented at full scale and sustained over a long period of time with political support. This was the DBE Workbooks series documented by McKay & Spaull. The Department of Basic Education has repeatedly credited this intervention as being one of the main likely explanations for the country's improvements in international assessments (e.g. DBE 2020). None of the other interventions, even those implemented on a relatively large scale, could have had a similar systemic impact. The main potential contribution of the other interventions lies in the extent to which they have generated evidence and learning to inform system-wide reforms. It is perhaps concerning then that most of the interventions described in this book were not designed with evaluation as a central activity, let alone to facilitate measurement of programme impact on learning outcomes. Yet, the fact that these interventions have been documented in this book does represent a shift in the sector towards valuing evaluation and learning from interventions. It is our impression that this level of concern to inform wider practice was not present to the same extent a decade ago.

3.3 Lessons about evaluation methods

Although there is now a concern to share learnings, only the various EGRS interventions, the Funda Wande interventions, and the Formative Assessment intervention (Kanjee & Bhana, this volume) were designed to measure causal impact. Several chapters apply statistical methods in an attempt to measure programme impact even though assignment to intervention group and data collections were not planned with evaluation in mind (for example Harris et al., this volume, Brombacher et al., this volume and Moloi et al., this volume). In these chapters, the ability to make valid causal inferences was limited. Where a more plausible method to identify causal estimates was well-executed, we have characterised those methods as quasi-experimental (Ardington & Spaull, this volume; Spencer-Smith et al., this volume) But where a more rudimentary approach was used, such as a simple pre- and post-test without a comparison group, we have referred to the statistics presented as 'descriptive' since they show how test scores changed over time but do not provide an estimate of the causal impact of the programme under review.

All the randomised control trials (RCTs) described in this book used a range of quantitative and qualitative methods to unpack the mechanisms underlying programme impact or lack thereof. Although they are good at measuring the ultimate causal effect of a programme, RCTs are sometimes criticised for not revealing why and how interventions worked (Deaton & Cartwright 2018). In the education context, the critique is that RCTs are designed to measure the impact on test scores but do not reveal the classroom dynamics that led to those improvements. But there is nothing about random assignment to intervention and control groups that prevents a range of qualitative evaluation methods being employed to better understand the intervention. Nor does a rigorous quantitative analysis prevent the use of qualitative methods. Random assignment and quantitative metrics of learning can and should be complemented by qualitative ethnographic and classroom observation studies, although it is true that they often are not. Yet the RCTs conducted through EGRS and by Funda Wande included in-depth lesson observation studies across 40 to 60 schools at a time, and conducted several rounds of case-studies in small numbers of schools (see Fleisch & Alsofrom). This has led to a range of complementary insights into these interventions. It also represents the collaboration of researchers coming from different traditions and with varying methodological specialisations.

3.4 Lessons about effect sizes

The chapters that used more robust methods to estimate causal effects of interventions (experimental or quasi-experimental methods) all tended to find small or moderate effect sizes. In contrast, those chapters using more rudimentary approaches to measure gains in test scores often suggested large impacts. EGRS had only one intervention with a moderate effect size (coaching), one intervention with a small positive effect (training), and three null results (a Grade 4 catch-up programme, a parental involvement programme, and a virtual coaching programme). This relationship between method and effect size is also apparent in a systematic review by Conn (2017). Across all pedagogy interventions reviewed the average reported effect on test scores was 0.92 standard deviations, but when the sample was restricted to methodologically high-quality studies the average effect size was a third of this (0.28 standard deviations). It would be perverse to respond to this observation by avoiding RCTs because they show small effects. Rather, it reveals that many if not most interventions that get implemented are having a negligible impact on learning and if they were rigorously evaluated they would be shown to be ineffective. Looking to the next decade we should subject more interventions to the discipline of rigorous impact evaluation, simply because without it claims about 'what works' are hollow and lack credibility.

In summary, we can distil the following lessons from the interventions described in this book:

1. There is compelling evidence that structured learning programmes using lesson plans, integrated reading materials, and professional support, and especially on-site coaching, can positively impact on reading.
2. The specific pedagogical design features of materials and interventions matter.
3. All successful interventions implemented to date provide considerably more materials to learners and teachers (anthologies of graded readers, lesson plans etc.) compared to business-as-usual schools.
4. Only the DBE Workbook programme is likely to have made a systemic impact. The primary value of the other interventions is to provide lessons for future system-wide reforms.
5. The most informative interventions were designed up-front to measure the causal effects of programmes, but also included a range of qualitative research methods.
6. Although there have been at least three mid-scale RCTs assessing reading interventions since 2015, there has been only one RCT on a mathematics intervention and only very recently, 2021 (Sapire et al.).
7. Effect sizes, when measured well, are usually small to moderate, implying that systemic improvement is not going to come easily, cheaply, or quickly.

4 Where to from here? The next ten years of interventions

Building on the lessons of the last decade, we have three recommendations for how to make more effective use of early grade reading and mathematics interventions in the next decade.

4.1 Return to the South African tradition of sustained system-wide reforms rather than piecemeal pockets of implementation

South Africa has a long tradition of implementing ambitious system-wide pro-poor reforms that have been sustained and subsequently credited with improvements in key outcomes. This includes the establishment and expansion of the Child Support

Grant, reaching 13.4 million children in 2021 (Van der Berg et al. 2022b), the roll-out of free antiretroviral treatment for HIV – the largest programme in the world – and in education the universal provision of free-school meals (JET, 2016). At the turn of the decade (2010) South African education policymakers were in keeping with this tradition, implementing a new curriculum, a new assessment system, and a new universal learner workbook programme (McKay & Spaull, this volume). In the last few years, however, there have not been any similar ambitious nationwide reforms in education. Instead, we now see a patchwork of well-intentioned projects and partnerships implemented in '25% of districts' or in '1,000 schools'. As a recent World Bank report on reading in South Africa puts it, "even though the national and provincial levels of government are either the leading partners or close contributors to the design and implementation of several successful interventions, these interventions are currently medium-scale pilots or experiments that are largely financed and implemented by external parties" (Kika et al. 2022, 42). While some smaller-scale experimentation is desirable, most of these 'pockets of implementation' do not sufficiently inform national policy. Too many interventions are neither at scale, nor informing scale. There is, in our view, a need to return to the South African tradition of building coalitions around evidence-based reforms that can be implemented in all schools.

We would argue that in 2022 there is now a broad consensus among policymakers, researchers and the public that early grade reading outcomes need to improve and this area should be prioritised. This is reflected in official government planning documents such as the Medium Term Strategic Framework (Presidency, 2019) and the DBE's Action Plan to 2030 (DBE 2020). The chapters in this volume provide a strong evidence base for the types of interventions that are likely to yield improvements in reading outcomes. The intervention should focus on the home-language (rather than EFAL) and include intensive in-classroom support to teachers that is aligned to a structured learning programme (lesson plans, workbooks and anthologies of graded reading stories).

The lessons of the DBE Workbook programme should inform this scale-up. It will require coordination across provinces and across the curriculum, LTSM and teacher-development sections within education departments. This kind of coordination is one of the key principles underlying the success of structured learning programmes: the materials provided, and the training or coaching is all directly targeted to facilitate the learning programme embodied in the lesson plans or workbooks. It is important to emphasise that the structured learning programme components are there to support implementation of the curriculum, rather than an alternative approach or something 'extra' to the curriculum. The implementation of this programme will need to be sustained over many years to create new norms of instructional practice. If teachers get used to the set of routines and materials used in this learning programme, then professional support in the form of training workshops and on-site coaching can occur on a rotational basis and serve to boost implementation quality. Although the country does not have the capacity or finance to provide intensive training workshops or on-site coaching to all primary schools at the same time, provinces could provide this to 20-30% of schools each year, and this could be a feasible practice if the same underlying learning programme and resources are expected to be used across the years. If the core components of this intervention (lesson plans and reading materials) are sustained over time and across the country, then coaching and training can be done on a rotational basis as part of a system-wide intervention. In contrast, if training and coaching projects are done in pockets of the system as part of separate projects, then these professional support initiatives will not have a systemic impact.

Even though the DBE Workbooks were not designed as a learning programme, but rather as a supplementary resource, they have fundamentally changed the routines of classroom practice. This presents an opportunity to leverage existing familiarity and 'buy-in' to the Workbook 'brand' to introduce revisions to the Workbooks and supplementary materials that incorporate some of the important pedagogic principles that other chapters in this book have described. Examples include the approach within the NumberSense and Bala Wande workbooks, the use of decodable texts as exemplified in the Vula Bula series, the approaches to formative assessment described in Kanjee & Bhana (this volume), and the use of mental mathematics used in the MSAP.

To maximise the take-up of any new system-wide interventions, these should be accompanied by refocusing the entire system on learning outcomes, in particular reading outcomes. As McKay & Spaull, (this volume) argue, one of the reasons for the success of the DBE Workbooks was that it was introduced at the same time as the Annual National Assessments, which although an imperfect intervention, did manage to draw attention to learning outcomes at the primary school level (Nuga Deliwe & Van der Berg, 2022). The president's (Ramaphosa, 2019) call for all children to learn to read for meaning by the age of ten is an easily communicated starting point. The DBE's recent work to develop reading benchmarks in all languages provides easily understood targets at the Grade 1, 2, and 3 levels (see Mohohlwane et al., 2022). For example, the fluency benchmarks of letters correct per minute (LCPM) and words correct per minute (WCPM) for the Nguni languages are 40 LCPM (Grade 1), 20 WCPM (Grade 2) and 35 WCPM (Grade 3). Although simple, these benchmarks were developed based on extensive analysis of the relationship between fluency and comprehension within this language family (Ardington et al. 2021). If children reach these critical benchmarks they can be regarded as on track to be reading with sufficient comprehension by the age of ten. It is now also well-established that fluency is a strong predictor of comprehension (Wills et al. 2022; Pretorius & Spaull 2022).

A sample-based reading survey could provide national and provincial estimates of the percentages of children reaching these critical benchmarks and learning to read for meaning by age ten. Teachers and subject advisors could also be trained to use existing tools like the Early Grade Reading Assessment. Combining new support interventions with a system that promotes accountability for learning can be regarded as the second wave of reforms, similar to the combination of workbooks, national assessments, and curriculum reforms around 2010, which seem to have contributed to South Africa's improvements in learning seen in the international assessments.

4.2 The need for rigorous evaluations of under-researched but strategically important points of intervention

Although there is a clear consensus that something should be done to improve early grade reading, it is not the only strategic point of intervention. There are at least four other strategic entry points that will increasingly become important:

1. With the migration of many Early Childhood Development (ECD) functions from departments of social development to the DBE, there is an opportunity to develop policies and programmes that drive quality improvement within more than 40,000 ECD programmes in South Africa. The case for the cost-effectiveness of early interventions applies even more strongly to ECD than to reading.
2. We need to build an evidence base around the options to improve early grade mathematics, although the positive results from the Bala Wande teacher assistant intervention is already providing some promising direction (Makaluza & Mpeta, this volume; Sapire et al., this volume; Ardington & Henry, 2021).
3. The children who were in Grades 1-3 during 2020 and 2021 and incurred huge learning losses as a result of the pandemics's disruption to schooling will now be moving through Grades 4 to 6. This presents an opportunity to develop catch-up programmes, possibly drawing on the so-called 'teaching at the right level' approaches that have delivered learning gains in places like India and Zambia (Banerjee et al. 2016; Vromant et al. 2021).
4. South Africa is currently entering a period of increased teacher retirements as the system's ageing workforce approaches retirement age. It is estimated that 45% of all publicly-employed teachers will retire in the next ten years (Van der Berg et al. 2020). This means that Initial Teacher Education becomes arguably the most strategic way to improve the quality of teaching across the system.

In order to generate robust evidence to inform system-wide interventions, those interventions not operating at scale should be designed to generate learning and evidence. We should move practice in the direction of the ideal phases of programme development and scale-up. Both Venkat & Graven (this volume) and Brombacher et al. (this volume) caution against going too quickly to the scale of an RCT (which requires at least about 40 intervention and 40 control schools to have enough statistical power) before doing careful design and piloting. A more common and worse problem, in our opinion, is going straight-to-scale before doing adequate design, piloting, or impact evaluation. Ideally, new interventions would go through several years of design and redesign, small pilots (fewer than 10 schools), then impact evaluation using experimental methods (80 - 300 schools) and finally scaling up to thousands of schools and the entire system. This is of course idealistic, but we do need to move further in this direction and take opportunities that may arise to influence for better design and evaluation. System-wide reforms should be put on hold until such time as a thorough evidence base has been developed, but we are arguing for a better use of research and evaluation as interventions are administered, especially in those not operating at full scale.

In the next decade we need to improve the way impact evaluations are planned and executed. It is of paramount importance that evaluation is planned, budgeted for, and influences the design of the intervention from the outset. Many of the complex statistical methods used at the analysis stage are needed to try (often unsuccessfully) to compensate for the lack of an evaluation design. To measure the causal impact of an intervention two fundamental things are needed: valid outcome measures, such as test scores, and a valid control group (providing a credible estimate of what outcomes would have been in the absence of the intervention). When a valid control group is not constructed up-front through random assignment of the intervention, it becomes complicated if not impossible to credibly estimate what the outcomes would have been without the intervention.

Aside from better evaluation design, we also need better norms in reporting on evaluation results so that readers can gauge the reliability of results and access all relevant information for scaling up decisions. Important information that should routinely be reported include how intervention schools were selected, how learners within schools were sampled, the sample sizes (of schools and of learners), what assessment instruments were used, the overall distributions and summary statistics of test scores, how many learners dropped out of the sample during the intervention period, and a breakdown of intervention costs per learner or per school.

A final recommendation for the next decade of evaluating interventions is to measure long-term impacts. The original EGRS interventions involved a three-year intervention in Grades 1 to 3 between 2015 and 2017. Subsequent learner testing was also conducted on the original sample of learners one year later (2018) and four years after the interventions ended (2021). The results indicate long-term positive impacts for those learners whose Grade 1 to 3 teachers had received lesson plans, integrated reading materials and coaching, but also document some fade-out of the effect – both important considerations when planning and designing subsequent interventions or scale-ups. In the next generation of interventions and evaluations, we need more measurement of long-term impacts, to understand which early-learning skills and intervention approaches are most impactful on later educational outcomes.

4.3 Avoid 'intervention fatigue' arising from a plethora of piecemeal interventions

Many school-support interventions, projects, and campaigns have been administered over the last decade, and many valuable lessons have been learned. However, one negative consequence of having many different interventions running across the system, especially when these are not sustained over time, is that teachers may become disillusioned about programmes they are asked to participate in. They are expected to comply with new approaches, and attend training sessions on weekends and holidays, only to be assigned to participate in something different the following year. We need to avoid creating this sort of intervention fatigue amongst teachers, and similarly avoid creating a type of intervention dependency at a provincial and district level. Provincial education departments need to be leading the implementation of a system-wide intervention and should allocate provincial budgets to the relevant materials and training. If provincial and national departments begin to rely on partnerships with a range of organisations running a variety of projects to plug gaps in early grade reading and mathematics, this will perpetuate intervention fatigue and intervention dependency, and will not lead to system-wide improvement.

5 Conclusion

Although there have been improvements over the past decade or more, foundational learning outcomes in South Africa are still unacceptably low and unequal. It has now been more than a decade since the important system-wide reforms introduced at the start of Minister Motshekga's first term (CAPS, Workbooks, ANA). Considering the well-documented learning losses of the pandemic (Ardington et al. 2021; Van der Berg et al. 2022), it is time for a new round of system-wide reforms that provide teachers with the kind of tools and support that has been proven to impact on learning through several of the interventions described in this book. This new round of reforms needs to marry effective teacher support with a strengthened accountability for learning and will require either additional budgets or re-allocation of budgets, both of which require political will and public support, especially in a time of austerity. The entire system, from government to teachers, parents, and learners, needs to be focused on the progress learners are making in achieving critical reading and mathematics outcomes in the first few grades.

The chapters in this volume illustrate quite clearly that the last decade of education interventions and evaluations has led to the creation of a large and local evidence base of 'what works'. Drawing on the national tradition of bold and systemic interventions in South Africa should embolden education policymakers to scale-up and sustain a few interventions that are sufficiently intensive and based on what is known about what actually changes teacher practice and learning outcomes. The next decade of interventions should be guided by greater clarity about their purpose and role. We need fewer interventions that are neither at scale nor generating evidence to inform scale. What we need is a new round of system-wide reforms that get sustained over many years, as well as several smaller interventions that are explicitly geared towards innovation, development, and evaluation to inform system planning. Piecemeal interventions or those that are large-scale but light-touch will not yield the improvements that are both direly needed and empirically possible.

References

Ardington, C., Wills, G., Pretorius, E., Mohohlwane, N., & Menendez, A. (2021). Benchmarking oral reading fluency in the early grades in Nguni languages. *International Journal of Educational Development*, 84, 102433.

Ardington, C., Wills, G., & Kotze, J. (2021). COVID-19 learning losses: Early grade reading in South Africa. *International Journal of Educational Development*, 86.

Balarin, M. (2021). The Political Economy of Education Reforms in Peru 1995-2020: Final Report.

Banerjee, A., Banerji, R., Berry, J., Duflo, E., Kannan, H., Mukherji, S., Shotland, M., & Walton, M. (2016). Mainstreaming an effective intervention: Evidence from randomized evaluations of "Teaching at the Right Level" in India. National Bureau of Economic Research. Working Paper No 22746.

Bruns, B., Evans, D., & Luque, J. (2012). Achieving World-Class Education in Brazil: The Next Agenda. Direction in Development; human development. World Bank. Available at: https://openknowledge.worldbank.org/handle/10986/2383 License: CC BY 3.0 IGO.

Bruns, B., Schneider, B., & Saavedra, J. (2022, Forthcoming). Technocrats and Turbulence: The Politics of Transforming Education in Peru: 2000-2020. Research on Improving Systems of Education (RISE).

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.

DBE. (2020). Action Plan to 2030. Department of Basic Education. (Online). Available at: https://www.education.gov.za/Portals/0/Documents/Publications/Sector%20plan%202019%2015%20Sep%202020.pdf?ver=2020-09-16-130709-860 (Accessed 14 November 2022).

Deaton, A., & Cartwright, N. (2018). Understanding and misunderstanding randomized controlled trials. *Social Science & Medicine*, 210, 2–21.

Evans, D. (2022). Every Child Reading: International Successes. Presentation at Early Grade Reading Research Indaba. Kiewietskroon, South Africa. 8 June 2022.

Gustafsson, M. (2019). How Fast Can Levels of Proficiency Improve? Examining Historical Trends to Inform SDG 4.1.1 Scenarios. UIS Information Paper No.62. Available at: http://uis.unesco.org/sites/default/files/documents/ip-62-how-fast-can-proficiency-levels-improve.pdf (Accessed 14 November 2022).

Human Sciences Research Council. (2012). Highlights from TIMSS 2011: The South Africa Perspective. Pretoria: HSRC.

JET Education Services. (2016). Report on the Implementation Evaluation of the National School Nutrition Programme. Available at: https://evaluations.dpme.gov.za/evaluations/528 (Accessed 14 November 2022).

Kika, J., Crouch, L., Ninan Dulvy, E., & Thulare, T. (2022). Early Grade Reading in South Africa. Policy Research Paper, World Bank Group.

London, J. D. (2021). Outlier Vietnam and the Problem of Embeddedness: Contributions to the Political Economy of Learning. RISE Working Paper Series. 21/062.

Mohohlwane, N., Ardington, C., & Wills, G. (2022). A review of recent efforts to benchmark early reading skills in South African languages. In Spaull, N. & Pretorius, E. (Eds) *Early Grade Reading in South Africa*. Cape Town: Oxford University Press.

National Planning Commission. (2012). National Development Plan 2030: Our Future-make it work. Available at: https://www.gov.za/sites/default/files/gcis_document/201409/ndp-2030-our-future-make-it-workr.pdf (Accessed 14 November 2022).

Nuga Deliwe, C., & Van der Berg, S. (2022). The role of assessment in Foundation Phase improvement: The Annual National Assessments and beyond. In Venkat, H. & Roberts, N. (Eds) *Early Grade Mathematics in South Africa*. Cape Town: Oxford University Press.

Pretorius, E., & Spaull, N. (2022). Reading research in South Africa (2010–2022): Coming of age and accounting for empirical regularities In Spaull, N. & Pretorius, E. (Eds) *Early Grade Reading in South Africa*. Cape Town: Oxford University Press.

Ramaphosa, C. (2019). President Cyril Ramaphosa: State of the Nation Address 2019. South African Government. Available at: https://www.gov.za/speeches/2SONA2019 (Accessed 14 November 2022).

Sandefur, J. (2022). Schooling for All Feasible Strategies to Achieve Universal Education. Washington D.C.: Center for Global Development.

Spaull, N. (2015). Accountability and capacity in South African education. *Education as Change*. Vol 19(3). 113–142.

Taylor, S., & von Fintel, M. (2016). Estimating the impact of language of instruction in South African primary schools: A fixed effects approach. *Economics of Education Review*, 50, 75–89.

Taylor, S., & Yu, D. (2009). The importance of socio-economic status in determining educational achievement in South Africa. Working Papers 01/2009, Stellenbosch University, Department of Economics.

Van der Berg, S., Gustafsson, M., & Burger, C. (2020). *School Teacher Supply and Demand in South Africa in 2019 and Beyond*. Pretoria: Department of Higher Education and Training.

Van der Berg, S., Patel, L., & Bridgman, G. (2022a). Food insecurity in South Africa: Evidence from NIDS-CRAM wave 5. Development Southern Africa.

Van der Berg, S., Hoadley, U., Galant, J., van Wyk, C., & Böhmer, B. (2022b). Learning Losses from COVID-19 in the Western Cape: Evidence from Systematic Tests. *Research on Socioeconomic Policy (RESEP)*. Stellenbosch: RESEP.

Vromant, N., Kuppens, L., Hazemba, M., & Kyulabantua, F. (2021). Scaling Teaching at the Right Level: Insights from the Catch-Up Literacy and Numeracy Intervention in Zambia. In Munene, I. (Ed.) *Ensuring All Children Learn: Lessons from the South on What Works in Equity and Inclusion*. Washington D.C.: Lexington Books.

Wills, G., Ardington, C., & Lesang Sebaeng, M. (2022). Foundational skills in home language reading in South Africa: Empirical evidence from 2015–2021. In Spaull, N. & Pretorius, E. (Eds) *Early Grade Reading in South Africa*. Cape Town: Oxford University Press.

index

Page numbers in italics indicate figures or tables

C

D

E

F

G

H

I

M

N

Q

S

T

Y

Z

EARLY GRADE READING AND MATHEMATICS INTERVENTIONS IN SOUTH AFRICA

INTERVENTIONS

The years since 2010 have yielded enormous evidence-based advances in our understanding of how South African learners successfully learn to read and do mathematics, and the interventions and materials that lead to those improvements.

This volume brings together 38 researchers from nine universities, various NGOs, and government to document how these interventions were conceptualised, implemented, and evaluated.

Whether reviewing the political factors surrounding the roll-out of the nationwide DBE Workbook programme, or the details of Randomised Control Trials and materials development programmes, all authors have provided detailed systematic accounts of how and why their interventions were structured and implemented as they were. The chapters included in this volume form arguably the largest and most up-to-date collection of early grade reading and mathematics interventions in South Africa. The volume includes research from across the full range of scale, including nationwide interventions (DBE Workbooks), medium-scale interventions (GPLMS, NECT, Vula Bula anthologies), and small-scale, localised interventions (including Bala Wande, EGRS, Funda Wande, JumpStart, NumberSense, Reading Eggs, R-Maths, and Room to Read).

The volume will interest local and international researchers, as well as policy-makers and donors, who want to know not only what works, but also why it works, how it works, and where it works.

“This volume makes it clear that South Africa has one of the richest collections of educational innovation, experimentation, and complementary research to evaluate reading and mathematics programs across the African continent. A decade of research is succinctly presented, documenting the country’s dynamic interplay between researchers, government, and non-government organizations. This volume will be crucial to policy-makers and researchers alike.”

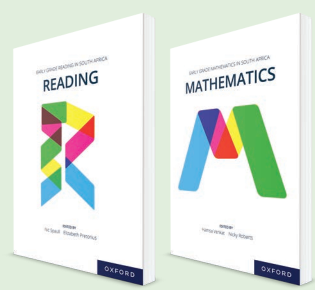
David K. Evans
Senior Fellow, Center for Global Development

EDITORS

Nic Spaull is an associate professor in the Research on Socioeconomic Policy (ReSEP) group within the Economics Department at Stellenbosch University.

Stephen Taylor is the Director of Research, Monitoring and Evaluation in the Department of Basic Education, South Africa.

Also available:



OXFORD
UNIVERSITY PRESS

www.oxford.co.za

A Proud Shareholder
THE MANDELA RHODES FOUNDATION

ISBN 978 0 19 074527 1

9 780190 745271

SOUTH AFRICA