LEADERSHIP FOR LITERACY

A Guide to Using Data from the 'Leadership for Literacy' Project

Name of study:

Leadership for literacy or officially "Succeeding Against the Odds: Understanding resilience and exceptionalism in high-functioning township and rural primary schools in South Africa".

Funder:
Economic and Social Research Council [grant ES/N01023X/1]

Principal Investigator:
Servaas van der Berg

Study period:
May 2016 - November 2018

Authors of this document:
Gabrielle Wills and Nicola Deghaye

Data submission: February 2019 (revised July 2019)

UNIVERSITEIT STELLENBOSCH UNIVERSITY

RESEP
Research on Socio-Economic Policy

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL
UKaid from the British people
Research jointly supported by the ESRC and DFID

Abbreviations used in the data:

EMS = Economic & Management Studies

FAL = First additional language

FP = Foundation phase

HOD = Head of department (within the school)

IP = Intermediate phase

LOLT = Language of learning & teaching

LSTM = Learner and Teacher Support Materials

ORF = Oral Reading Fluency

SLM = School leadership and management

SMT = School management team. SMT includes HODs, deputies and principal.

SGB = School governing body

Contents

List of tables

1. Introduction

This document provides a guide for users interested in the quantitative and case study data collected for the "Leadership for Literacy" project in 61 South African schools. This document first provides a brief overview of the project and guides the reader to research outputs identifying project findings. It then explains what data was collected and provides specific details about variables collected in individual datasets. This user guide should be read alongside other research outputs related to the project to gain a better understanding of the objectives and intents of the project. Additionally, the various codebooks of meta-data and instruments related to individual datasets collected are identified in this document.

1.1 About the project

"Understanding resilience and exceptionalism in high-functioning township and rural primary schools in South Africa", more affectionately known as "Leadership for literacy", is an education research project lead by a multi-disciplinary team of researchers across Stellenbosch University, the University of Cape Town, JET, the University of South Africa (UNISA) and the Department of Basic Education (DBE) in South Africa. The project was initiated in reaction to a deficit discourse where much research has focussed on the realities of a highly underperforming schooling system in South Africa. Where solutions are desperately needed in the no-fee sector, less consideration has been given to exploring pockets of higher performance that may exist.

We briefly identify the five broader objectives that motivated the study in the table below, while referring the reader to working papers which explore the methodological approaches and findings associated with each of the four objectives.

Table 1: Research objectives and related reference documents

Research objective	Relevant research output
1) Identify the number of exceptional rural and township primary schools in South Africa. *We aimed to establish how many exceptional schools, catering to the poor (i.e. that are fee-free or no-fee charging), exist in three provinces in South Africa.* *This search process is of importance, as information on exceptional schools in the country has been anecdotal, or (for secondary schools) based on good performance in a single evaluation or matriculation examination.* *The results of this search however lead to very disappointing conclusions about the absence of high-quality schools for the poor in three South African provinces.*	Wills, G. (2017) What do you mean by 'good'? The search for exceptional primary schools in South Africa's no-fee school system. *Stellenbosch Working Paper Series No. WP16/2017. Stellenbosch University.* https://www.ekon.sun.ac.za/wpapers/2017/wp162017
2) Gaining new insights into school leadership and management practices in high achieving schools relative to average or low-achieving schools in challenging contexts. *This involved a strong mixed-methods research design with qualitative case studies and quantitative data collection informing this research process. Eight school case studies were*	Hoadley, A (2018) Leading for literacy: A review of the research. RESEP working paper. Stellenbosch University. http://resep.sun.ac.za/leading-for-literacy-a-review-of-the-research/ Taylor, N. and Hoadley, U. (2018) Leadership for Literacy: Exploring

used to generate thick descriptions of school leadership and management (SLM) processes and practices, and was supported through quantitative data collection using a new set of instruments developed for the study.	leadership practices in township and rural primary schools serving poor communities in South Africa. Final Report on the Case Study Schools. *RESEP Report.* http://resep.sun.ac.za/leadership-for-literacy-final-report-on-the-case-study-schools/ See also Taylor, Wills and Hoadley (2018) below.
3) The development of a new School Leadership and Management (SLM) instrument that captures the practices and behaviours of teachers and principals in challenging contexts in South Africa. *We aimed to develop a quantitative tool to capture leadership and management practices in South Africa that are theoretically linked to literacy. The intention was that the design of this tool support an 'at scale' application.* *The instrument development process was strongly underpinned by a theoretical framework derived for this study:*	Taylor, N, Wills, G. and Hoadley, U. (2018) Addressing the 'leadership conundrum' through a mixed methods study of school leadership for literacy. RESEP working paper. Stellenbosch University. http://resep.sun.ac.za/addressing-the-leadership-conundrum-through-a-mixed-methods-study-of-school-leadership-for-literacy/
4) To determine the predictive validity of the developed SLM instrument. *We set out to determine whether our quantified measures of SLM were predictive of both i) measures of grade 6 and grade 3 learner literacy and reading levels and ii) intermediate outcomes such as curriculum coverage, teacher absenteeism and teacher motivation.*	Wills, G. and van der Berg, S. (2018) Measuring leadership and management and their linkages with literacy in rural and township primary schools in South Africa. *Stellenbosch Working Paper Series No. WP21/2018.* *https://www.ekon.sun.ac.za/wpapers/2018/wp212018* See also Taylor, Wills and Hoadley (2018) above.
5) To establish tentative early grade reading norms and benchmarks among South African learners in challenging contexts. *There has been an increased discourse that attaches value to African languages in South Africa, and the importance of home language instruction. Yet, the home language literacy proficiencies of learners are substantially below expected international standards and we know little about how reading is being taught in African languages in South Africa. The Progress in International Literacy and Reading Study (2016) highlighted that by the end of grade 4, 78% of learners cannot read for meaning in any language. In this context, we developed reading tests in three of 11 official African languages. The data collected can be used to establish tentative norms and reading benchmarks. It also provides*	Spaull, N., Pretorius, L. and Mohohlwane, N. (2018) Investigating the Comprehension Iceberg: Developing empirical benchmarks for early grade reading in agglutinating African languages. *RESEP working paper.* *https://www.jet.org.za/clearinghouse/primted/resources/language-and-literacy-resources-repository/wp-v11-esrc-paper-1-comprehension-iceberg-v4.pdf/@@images/a7b5b7c2-95b8-4b78-b48b-6c2efbf15de9.jpeg* Wills, G. and Hofmeyr, H. (2018) Academic Resilience in Challenging

outcome measures to assess the predictive validity of our SLM measures (objectives 3 and 4).	Contexts: Evidence from Township and Rural Primary Schools in South Africa. *Stellenbosch Working Paper Series No. WP18/2018.* https://www.ekon.sun.ac.za/wpapers/2018/wp182018

Related available policy briefs:

Creating fluency benchmarks in African languages (December 2018)
by Nic Spaull, Elizabeth Pretorius and Nompumelelo Mohohlwane
http://resep.sun.ac.za/creating-reading-fluency-benchmarks-in-african-languages/

Exploring how school leaders promote literacy improvements (December 2018)
by Nick Taylor, Gabrielle Wills and Ursula Hoadley
http://resep.sun.ac.za/exploring-how-school-leaders-promote-literacy-improvements/

Academic resilience in challenging school contexts in South Africa (December 2018)
By Gabrielle Wills and Heleen Hofmeyr
http://resep.sun.ac.za/academic-resilience-in-challenging-school-contexts-in-south-africa/

The possibility of improvements despite a lack of high-quality no-fee primary schools (December 2018)
By Gabrielle Wills
http://resep.sun.ac.za/the-possibility-of-improvements-despite-a-lack-of-existing-high-quality-no-fee-primary-schools/

2. Ethics

Throughout this project we have followed ethical processes in obtaining permissions to conduct research, in obtaining consent for participation and maintaining the anonymity of respondents.

At project inception we received written permission to conduct fieldwork in schools from each of the provincial education departments in the three provinces in which we would conduct research. We also received written approval from the Director General of the National Department of Basic Education, supporting the project, supporting research in schools, and providing permission to use the Annual National Assessment data.

Ethical clearance for the research was granted by the Stellenbosch University Research Ethics Committee, which also involved the approval of administered instruments.

All interviews and learner testing commenced with obtaining consent from research participants. Additionally, opt-out forms were sent to schools prior to learner testing. If parents returned an opt-out letter via the learner, then learners did not participate in the testing.

During testing, learner names were written on separate tear-off pages from assessment booklets to protect the anonymity of the learners and their test results.

Finally, we have taken care to remove the names of schools and teachers from the quantitative data and qualitative case-study data.

3. School Sampling

The first objective of the project, identifying the number of exceptional rural and township primary schools in South Africa, motivated the use of a purposeful school sampling approach.

We engaged in an intensive search to identify the 30 best possible high-performing no-fee schools in three provinces (**KwaZulu-Natal, Gauteng and Limpopo**). The three provinces are chosen to represent distinct levels

of administrative functionality: Gauteng (a highly functional administration), KwaZulu-Natal (medium functionality) and Limpopo (low functionality).

This search process (described in detail in Wills (2017))[1] relied on identifying top performing schools in system-wide low stakes testing data – namely the Annual National Assessments - corroborated against a large dataset we collected of recommended 'good' schools from a host of sources (district officials, school principals and administrative clerks, education related NGOs, unions, other stakeholders, secondary schools performing well in the school-leaving[2] certificate). Then 30 lower performing schools located in similar geographic locations as the higher performing pairs were included in the sample. The schools were also selected based on either their language of instruction or the dominant student language in the school being **Zulu, Sepedi or Xitsonga.** (Please see Wills 2017 for a fuller discussion of why each of the schools was selected for the sample).

At project onset, a mixed methods approach with a matched pairs design was envisaged. Each outlier school is paired with a nearby typical or underperforming school. The matched pairs approach assumes that given a similar geographical position each school should share similar socioeconomic characteristics and cultural/political/local dynamics. This largely supports the qualitative component of the project. By making comparisons across high-performing and low-performing schools one can factor out the influence of some unobserved characteristic on the findings. However, the first challenge in establishing a set of schools to visit was to *identify* the outlier school pairs. This intention informed the **purposive school sampling approach.**

The purposeful sampling approach was also hypothesised to aid the detection of leadership effects, in estimating the relationship between management and leadership indicators collected in our study and literacy outcomes. A possible reason cited in the international literature for the identification of weak associations between school management and leadership and student learning is that study samples selected lack variation in both student learning and SLM practices. The 60 schools were purposively selected to artificially add as much student performance variation as possibly exists in the available sampling frame of public schools reaching poorer student populations in three provinces.

It must be noted that in February/March we surveyed 61 schools, however 1 school refused to be surveyed in October, leading to a final pre- and post-test sample of 60 schools.

4. Literacy assessments and student sampling

Pre-test literacy and reading scores for grade 3 and grade 6 learners were obtained at the beginning (February/March) of the 2017 year in 61 schools. This process was necessary to verify the quality of the selected schools and to select case-study schools for the qualitative work. The same tests were administered again at the end of year to (October/November) of the same year to obtain post-test scores for students. In this respect we have a panel of assessment data for grade 3 and grade 6 learners.

4.1 Grade 6 literacy data

The grade 6 English literacy test consisted of a written silent reading comprehension test and written vocabulary test administered to **one entire class of grade 6 students in each school (unless opt-out forms were returned by the learners to the school)**. Of the original pre-test sample of 2 656 students, 2 379 wrote the post-test, indicating a low attrition rate of 11%. The two comprehension tests consisted of released

[1] Wills, G., 2017. What do you mean by 'good'? The search for exceptional primary schools in South Africa' s no-fee school system (No. WP16/2017), Stellenbosch Economics Working Paper Series.

[2] The National Senior Certificate or otherwise known as matriculation examination.

items from previous rounds of the grade 4 PIRLS assessment. Permission was received from the IEA for their use. The reliability of these assessments is reflected in a high correlation between pre-test and post-test scores[3].

Additionally, pre- and post-test English one-on-one assessments including Oral Reading Fluency (ORF) are available for 599 grade 6 students. The grade 6 learners tested one-on-one are a sub-sample of the classes completing the written assessment. **Specifically, fieldworkers administered oral one-on-one assessments to 2 of the grade 6 teachers' best identified learners and then a random sample of an additional 8 to 10 students from the same class list.** ORF tests in African language were also administered to these students but only at the end of the year.

All one-on-one assessments were administered on tablets using Tangerine software, developed by RTI. The assessments were typically conducted in a quiet room to prevent distractions during the assessment.

4.2 Grade 3 literacy data

At the grade 3 level, a battery of literacy and reading tests – in English and African language (Sepedi, Xitsonga or Sepedi) - were administered. Pre- and post-test scores are available for 631 grade 3 students. The random selection of the Grade 3 students within each school for one-on-one assessments was executed as follows (and follows the same procedure as for the grade 6 one-one-one assessment sampling). First, the teacher was asked to select her top two learners. The fieldworker then randomly selected from the class list a further 8 or more learners to test. In some schools, the fieldworkers were not able to complete the full number of learner assessments as required.

4.3 Details on the literacy and reading assessments

The following table provides details on the filenames of the relevant literacy datasets, related assessment instrument booklets and the codebook meta-data for individual variables in each dataset. The pre- and post-test scores are linked for each student so that there is only one grade 3 learner assessment dataset and one grade 6 learner assessment dataset despite two 'waves' of testing. In the case of the grade 6 data, one-on-one assessment data for a sub-sample of those that wrote the written assessment is linked to learners' written assessment data in the same dataset.

Table 2: Learner assessments - a guide to dataset filenames, instruments and codebooks

	Dataset (filename)	**Instrument (filename)**	**Codebook (filename)**
Grade 3 learner assessment data	gr3_feb_oct_clean.dta	gr3_test_instruments_all.pdf	codebook_gr3_test.pdf
Grade 6 learner assessment data	gr6_feb_oct_clean.dta	gr6_oraltest_instrument_all.pdf gr6_writtentest_feb_instrument.pdf gr6_writtentest_oct_instrument.pdf gr6_writtentest_marking_guide.pdf	codebook_gr6_test.pdf

In the tables that follow we provide more information on the reading and literacy assessments.

- Table 3 provides a summary of learner sample sizes for grades 6 and 3 by language of assessment.
- Table 4 summarises the components of the assessments i.e. what they include.
- Table 5 and 6 provides a summary of descriptive statistics of grade 6 and grade 3 learner performance on the assessments.

[3] The Pearson correlation statistic between the pre-test and post-test was almost 0.90.

Table 3: Learner sample sizes

	English			Zulu (Gauteng & KwaZulu-Natal)			Sepedi (Limpopo)			Xitsonga (Limpopo)		
	Pre-test	**Post-test**	**Pre- & post-test**	**Pre-test**	**Post-test**	**Pre- & post-test**	**Pre-test**	**Post-test**	**Pre- and post-test**	**Pre-test**	**Post-test**	**Pre- and post-test**
Schools	61	60	60	42	42	42	8	8	8[c]	10	10	10
Learners:												
Grade 3 one on one assessments[a]	756	633	633	509	429	429	135	114[c]	114	111	89	89
Grade 6 one on one assessments[b]	733	599	599	NA	386[d]	NA	NA	93	NA	NA	110	NA
Grade 6 written assessments	2656	2544	2382	NA	NA	NA	NA	NA	NA	NA	NA	NA

Notes:

[a] If a grade 3 learner was assessed in English, they were also assessed in an African language

[b] Grade 6 learners who were assessed on-on-one comprise a sub-sample of a class who wrote the written literacy test. In February (i.e. pre-test) only English one-on-one assessment was conducted with Grade 6 learners. In October, English and African language assessments were conducted with Grade 6 learners.

[c] Only 1 grade 6 learner completed both the pre- and post-test in schoolid = 4300. Mixing of LOLT across classes prevented testing the same learners twice.

[d] There are very peculiar outcomes for schoolid = 2400 with respect to African (isiZulu) grade 6 ORF results. There were 10 post-test observations from this school that should be excluded in the analysis of grade 6 African language results.

Table 4: Learner assessments administered

Skill	**Sub-task & measurement**	**Pre-test (beginning of school year)**		**Post-test (end of school year)**	
		isiZulu/Sepedi /Xitsonga	**English**	**isiZulu/Sepedi/ Xitsonga**	**English**
Letter recognition	Number of letters read correctly in 60 seconds	Grade 3		Grade 3	
Word Recognition	Familiar word reading, number of correct words read in 60 seconds	Grade 3	Grade 3	Grade 3	Grade 3
Non-word recognition	Non-word reading, number of correct words read in 60 seconds	Grade 3		Grade 3	
Oral Reading Fluency	Oral passage reading, number of words fluently read (with accuracy) from a reading passage in 60 seconds	Grade 3 only (Same passage translated into different languages to establish benchmarking norms)	Grade 3 and 6	Grade 3 ***& 6*** (Same passage translated into different languages to establish benchmarking norms)	Grade 3 and 6
Reading compre-hension	Questions answered about the passage read aloud by the student. Questions are only asked up to a point at which the stopped student reading.	Grade 3	Grade 6	Grade 3 **& 6**	Grade 3 and 6
Written vocabulary test	Complete sentence, filling in the missing word. Section A tests 2000 most common words in English.		Grade 6 & *Grade 6 teacher*		Grade 6
Written compre-hension	Number of questions answered correctly about two PIRLS written passage in a pen and paper assessment		Grade 6		Grade 6

Table 5: Summary statistics for Grade 6 learner assessment data

Grade 6 results	Units of reporting	Mean	SD	P10	P25	P50	P90	Min. Mark	Max. Mark	Max. possible mark	N	N with negative gains
English Literacy (Written)												
Written Pre-test	*total marks*	13.1	9.9	4	6	10	27	0	63.5	106	2652	
Written post-test	*total marks*	18.0	12.9	5.5	8.5	14.5	37	0	75	106	2541	
Change on written test	*total marks*	5.1	6.1	-1.5	1	4.5	13	-36.5	36		2379	404
Change on written test	*std. deviation*	0.5	0.6	-0.2	0.1	0.5	1.3	-3.7	3.6		2379	404
PIRLS text 1 pre-test	*total marks*	4.4	3.1	1	2	4	9	0	16	16	2652	
PIRLS text 1 post-test	*total marks*	5.2	3.5	1	2	4	10.5	0	16	16	2541	
Change on PIRLS text 1	*total marks*	0.8	2.6	-2	-1	1	4	-9	11		2379	755
PIRLS text 2 pre-test	*total marks*	3.5	2.4	1	2	3	7	0	15	18	2652	
PIRLS text 2 post-test	*total marks*	4.2	2.8	1	2	4	8	0	16	18	2541	
Change on PIRLS text 2	*total marks*	0.7	2.4	-2	-1	1	4	-11.5	10.5		2379	776
Total vocabulary pre-test	*total marks*	5.1	5.7	0	1	3	13	0	42.5	72	2652	
Total vocabulary post-test	*total marks*	8.6	7.9	1	3	6.5	19	0	53	72	2541	
Change on vocabulary	*total marks*	3.6	4.1	-0.5	1	3	9	-26.5	31.5		2379	255
Vocabulary A pre-test	*total marks*	2.5	2.4	0	0.5	2	6	0	12.5	18	2652	
Vocabulary A post-test	*total marks*	3.9	3.0	0.5	1.5	3.5	8	0	15.5	18	2541	
Change on vocabulary A	*total marks*	1.5	1.8	-0.5	0	1.5	4	-8.5	9.5		2379	323
English Oral Reading Fluency												
ORF English pre-test	*total WCPM*	81.1	38.4	28	59	80	130	0	202	203	733	
ORF English post-test	*total WCPM*	93.4	40.0	40	68	97	144	0	192	203	599	
Change in ORF English	*total WCPM*	12.7	17.4	-5	3	12	32	-95	92		598	104
Change in ORF English	*std. deviation*	0.3	0.5	-0.1	0.1	0.3	0.8	-5.3	2.4		599	105
ORF English pre-test	*% WCPM*	40.0	18.9	13.8	29.1	39.4	64.0	0	99.5		733	
ORF English post-test	*% WCPM*	46.0	19.7	19.7	33.5	47.8	70.9	0	94.6		599	
ORF Sepedi post-test	*% WCPM*	19.6	9.9	6.2	11.7	21.1	31.2	0.6	45.1		93	
	total WCPM	60.5	30.6	19	36	65	96	2	139	308	93	

ORF Xitsonga post-test	*% WCPM*	23.1	12.0	6.9	15.6	22.7	36.8	0	55.2		110	
	total WCPM	66.6	34.4	20	45	65.5	106	0	159	288	110	
ORF Zulu post-test	*% WCPM*	26.0	10.5	13.9	17.5	26.5	38.0	0	56.6		386	
	total WCPM	43.1	17.5	23	29	44	63	0	94	166	386	
ORF combined African post-test	*% WCPM*	24.4	11.0	11.1	16.9	24.7	36.8	0	56.6		589	

Table 6: Summary statistics for Grade 3 learner assessment data

Grade 3 Results	Units	Mean	SD	P10	P25	P50	P90	Min. mark	Max. mark	Max possible mark	N	N with negative gains
English Results												
English ORF pre-test	%WCPM	20.8	16.6	0.0	6.3	20.5	41.7	0.0	85.8		755	
	total WCPM	22.2	24.8	0.0	0.0	15.0	59.0	0.0	127.0	127	731	
English ORF post-test	%WCPM	31.2	20.4	0.0	17.3	29.9	59.1	0.0	98.4		632	
	total WCPM	39.7	33.2	0.0	11.5	34.0	87.0	0.0	127.0	127	632	
Change in English ORF	total WCPM	17.6	17.3	0.0	3.0	16.0	40.0	-53.0	127.0		611	27
	standard deviation	0.7	0.7	0.0	0.1	0.6	1.6	-2.1	5.1		611	27
SWRC pre-test	%WCPM	21.2	18.8	0.0	4.5	18.2	46.6	0.0	93.2		731	
	total WCPM	18.7	16.5	0.0	4.0	16.0	41.0	0.0	82.0	88	731	
SWRC post-test	%WCPM	34.3	23.2	2.3	17.0	31.8	65.9	0.0	100.0		632	
	total WCPM	30.2	20.4	2.0	15.0	28.0	58.0	0.0	88.0	88	632	
Change in SWRC	%WCPM	13.2	12.7	0.0	4.5	12.5	27.3	-85.2	97.7		611	39
Zulu Results												
ORF pre-test	%WCPM	30.7	22.8	0.0	8.8	30.9	61.8	0.0	98.5		509	
	total WCPM	20.9	15.5	0.0	6.0	21.0	42.0	0.0	67.0	68	509	
ORF post-test	%WCPM	46.3	28.2	0.0	27.9	48.5	83.8	0.0	100.0		429	
	total WCPM	31.5	19.2	0.0	19.0	33.0	57.0	0.0	68.0	68	429	
Sepedi Results												
ORF pre-test	%WCPM	29.0	20.5	0.0	8.8	32.8	53.6	0.0	87.2		135	
	total WCPM	36.2	25.7	0.0	11.0	41.0	67.0	0.0	109.0	94	135	
ORF post-test	%WCPM	44.0	23.9	1.6	31.2	47.2	68.8	0.0	100.0		113	

	total WCPM	55.0	29.9	2.0	39.0	59.0	86.0	0.0	125.0	94	113	
Xitsonga Results												
ORF pre-test	%WCPM	37.9	24.7	0.0	12.4	44.8	67.6	0.0	86.7		111	
	total WCPM	39.8	25.9	0.0	13.0	47.0	71.0	0.0	91.0	125	111	
ORF post-test	%WCPM	56.5	28.3	0.0	42.9	59.0	89.5	0.0	100.0		89	
	total WCPM	59.4	29.7	0.0	45.0	62.0	94.0	0.0	105.0	125	89	
Combined African lang. results												
ORF pre-test	%WCPM	31.4	22.9	0.0	10.3	33.3	59.2	0.0	98.5		755	
ORF post-test	%WCPM	47.4	27.7	0.0	29.4	49.8	83.8	0.0	100.0		632	
Change in ORF post-test	%WCPM	15.5	14.4	0.0	5.9	14.7	32.4	-75.0	100.0		631	18
SWRC pre-test	%WCPM	34.2	22.0	0.0	15.0	36.0	63.3	0.0	88.0		755	
SWRC post-test	%WCPM	49.3	24.7	10.0	36.0	52.0	80.0	0.0	100.0		632	
Change SWRC	%WCPM	14.5	12.7	0.0	6.0	14.0	28.3	-78.0	93.3		631	21

Notes: SWRC = single words read correctly. WCPM = words correct per minute. ORF = Oral reading fluency.

4.1 Details on the assessments and marking information

Marking of the one-on-one literacy assessments is an automatically data captured process that occurs on Tangerine as the child is assessed. Some of the specific tasks are timed assessments as identified in the table below. This must be considered when evaluating the missing data in the text. Missing data could reflect that the child was not assessed, or a specific word/letter was not assessed but the child was assessed. In the data, item level marks are provided for each individual assessment – i.e. whether exact letters, or words were read correctly or incorrectly.

The grade 6 English written assessment at pre-test and then post-test was marked in office. Each marker was provided the marking memo as seen in "gr6_writtentest_marking_guide.pdf". A random sample 10% of all scripts was remarked by an expert marker. If consistent problems were found, the respective marker was asked to check his/her scripts again. Marks were double captured to detect errors in marking. Item level marks are provided for all reading comprehension questions, however, the vocabulary test marks are entered as totals for each section (A, B, C, D).

Table 7: Grade 3 assessments (pre-test and post-test are the same)

1. Grade 3 African language Oral Reading Fluency assessment (isiZulu, Sepedi or Xitsonga)		
Task	**Timed task**	**Description**
TASK 1: Letter recognition *isiZulu max. letters: 110* *Xitsonga max. letters: 110* *Sepedi max. letters: 110*	Yes (1 minute)	Letters were assessed as well as digraphs and trigraphs. The assessor identifies the specific letters/digraphs/trigraphs the learner reads correctly (1) and incorrectly (0).
TASK 2: Familiar word recognition *isiZulu max. words: 50* *Xitsonga max. words: 60* *Sepedi max. letters: 60*	Yes (1 minute)	The assessor identifies the specific words the learner reads correctly (1) and incorrectly (0).
TASK 3: Non-word recognition *isiZulu max. words: 50* *Xitsonga max. words: 50* *Sepedi max. letters: 50*	Yes (1 minute)	The assessor identifies the specific words the learner reads correctly (1) and incorrectly (0).
TASK 4: Oral reading fluency *isiZulu max. words: 68* *Xitsonga max words: 94* *Sepedi max. letters: 125*	Yes (1 minute)	Children were asked to read from the passage "Isobho Lamatshe" / "Sopo ya maribye"/ "Sopo ya maswika". The assessor identifies the specific words the learners gets correct (1) and incorrect (0).
TASK 5: Comprehension questions (asked orally)	No	See page 42 / 73 / 103 of the gr3_test_instrument_all for an indication of ORF word cut-offs associated with each comprehension question in isiZulu / Xitsonga / Sepedi.
2. Grade 3 English Oral Reading Fluency assessment		
Task	**Timed task**	**Description**
TASK 1: Assessment of familiar words (word recognition) *Max letters: 88*	Yes (1 minute)	The assessor identifies the specific words the learner reads correctly (1) and incorrectly (0).
TASK 2: Oral reading fluency PART 1 "Can the learner read the title of the story?" and "What is the story about?"	No	This is a preliminary question to ascertain whether the learner can read at all. If they cannot read one word of the title the rest of the assessment should have been discontinued.
TASK 2: Oral reading fluency PART 2 Reading the passage *Max. words: 127*	Yes (1 minute)	Children were asked to read from the passage "How the elephant got its trunk". The assessor identifies the specific words the learners gets correct (1) and incorrect (0).
TASK 4: Comprehension questions (asked orally)	No	There are 9 comprehension questions which relate the ORF passage "How the elephant got its trunk". Importantly, the number of questions asked of the child depends on how far they read in the ORF passage. See page 16 of the gr3_test_instrument_all for an indication of ORF word cut-offs associated with each comprehension question.

Table 8: Grade 6 assessments (pre-test and post-test are the same)

1. Grade 6 English language Oral Reading Fluency assessment		
Task	**Timed task**	**Description**
TASK 1: Oral reading fluency *Max. words: 203*	Yes (1 minute)	Children were asked to read from the passage "An unbelievable night". The assessor identifies the specific words the learners gets correct (1) and incorrect (0).
TASK 2: Comprehension questions (asked orally)	No	There are 8 comprehension questions which relate the ORF passage "An unbelievable night". Importantly, the number of questions asked of the child depends on how far they read in the ORF passage. See page 12 of the gr6_oraltest_instruments_all for an indication of ORF word cut-offs associated with each comprehension question.
2. Grade 6 African language Oral Reading Fluency assessment (isiZulu, Xitsonga, Sepedi) Note: This assessment was only done at Post-test (October/November 2017)		
Task	**Timed task**	**Description**
TASK 1: Oral reading fluency *isiZulu max. words: 166* *Xitsonga max words: 288* *Sepedi max. letters: 308*	Yes (1 minute)	Children were asked to read from the passage "Iyanyonyoba Inyosi Entinyelayo "/ "Mi va na vuxiyaxiya ku suka endzhaku"/ "Hlokomelaaa-ng ka moraaa-go". The assessor identifies the specific words the learners gets correct (1) and incorrect (0).
TASK 2: Comprehension questions (asked orally)	No	There are 10 comprehension questions which relate the ORF passage. Importantly, the number of questions asked of the child depends on how far they read in the ORF passage. See page 12 / 29 / 38 of the gr6_oraltest_instruments_all for an indication of ORF word cut-offs associated with each comprehension question in in isiZulu / Xitsonga / Sepedi.
3. Grade 6 English written assessment		
Task	**Timed task**	**Description**
Written comprehension 1: Fly, Eagle, Fly *Max marks: 16*	Yes (40 minutes – 20 minute break after this test)	Fly Eagle Fly is a released passage from PIRLS. Permission was obtained from the IEA and the author of the passage for its use. Please see the marking memo (gr6_writtentest_marking_guide.pdf) for an indication of marks assigned to each question item.
Written comprehension 2: The Giant Tooth Mystery *Max marks: 18*	Yes (40 minutes – 20 minute break after this test)	The Giant Tooth Mystery is a released passage from PIRLS. Permission was obtained from the IEA and the author of the passage for its use. Please see the marking memo (gr6_writtentest_marking_guide.pdf) for an indication of marks assigned to each question item.
English vocabulary test: *Max marks: 72*	Yes (40 minutes)	Section A tests knowledge of the 2000 most common words in English. Section B, C, and D test increasingly less common words in English. Half marks were given if the learners phonetically indicated that they knew the word but the spelling was incorrect.

5. Case Studies

As part of the broader mixed methods study, a key objective was to gain new insights into school leadership and management practices in high achieving schools relative to average or low-achieving schools in challenging contexts. The selection of 8 case study schools was central to achieving this objective. A link to the synthesis report on the case study findings is provided in:

> Taylor, N. and Hoadley, U. (2018) Leadership for Literacy: Exploring leadership practices in township and rural primary schools serving poor communities in South Africa. Final Report on the Case Study Schools. *RESEP Report.*
>
> http://resep.sun.ac.za/leadership-for-literacy-final-report-on-the-case-study-schools/

However, a wider team of qualitative and quantitative specialists completed individual case study reports for the selected 8 schools. These reports are anonymised and released with the following filenames as reflected in the table below. School pairs are represented by the letters A, B, C, D while the suffix H indicates that the school was the higher performing pair and L indicates that it was a lower performing pair.

Table 9: List of case study data documents -guide to filenames

Filename	Document name
AH_case_study.pdf	Case Study School A(H): Raw case study notes
AL_case_study.pdf	Case Study School A(L): Raw case study notes
BH_case_study.pdf	Case Study School B(H): Raw case study notes
BL_case_study.pdf	Case Study School B(L): Raw case study notes
CH_case_study.pdf	Case Study School C(H): Raw case study notes
CL_case_study.pdf	Case Study School C(L): Raw case study notes
DH_case_study.pdf	Case Study School D(H): Raw case study notes
DL_case_study.pdf	Case Study School D(L): Raw case study notes

5.1 Sampling

A matched pairs design guided the selection of case study schools. Four higher achieving schools were selected from the larger 60 school sample if they exhibited higher grade 6 literacy performance scores when measured at the medium and this performance exceeded expectations in terms of the socio-economic status level of the learners in the grade 6 sampled class.

Two main factors determined the choice of lower performing matched pairs:

1) Performance of the matched school in grade 6 literacy and at least one grade 3 language (English or African language) must be sufficiently below the performance of the outlier, again using the median class performance as the measure.

2) There had to be enough overlap in the socio-economic status of the grade 6 class tested with the socio-economic status of grade 6 learners in the higher achieving pair.

The matched pair preferably was chosen to have lower grade 3 performance in African and English language but this was not always possible.

Note: The final matched pairs selected are NOT the same as the original matching pairs. We had to rematch for the following reasons:

i) ANA was not a useful measure for the original matching purpose. It was a weak metric for our purposes with many schools that were outliers on ANA underperforming relative to their under-performing ANA matched pair when we analysed performance in our grade 6 literacy test.

ii) Neighbouring schools do not necessarily serve students of a similar SES. This is a fundamental problem with usual matching approaches adopted in qualitative work which became evident through our survey data where we collected asset information from students. We re-matched on socio-economic status using our index of asset ownership identified by the grade 6 students that were tested in each class. This is by no means a perfect metric of SES but certainly better than assuming that nearby schools serve similarly wealthy or poor students.

6. An overview of the contextual and school leadership and management related instruments

In addition to the learner assessment data, a host of instruments were administered to collect contextual and background information on schools, students and teachers.

Many of these instruments were designed with the intention of collecting information to quantify school leadership and management processes and practices in schools. The finalised SLM questions, which were distributed across 6 different instruments, were only administered in October. The majority of the February/March instruments were treated as pilot instruments used to explore whether questions provided any useful management information. The questions were completely changed for October/November data collection, with closer alignment to our "leadership for literacy theoretical framework" so that most variables related to these other instruments are not 'longitudinal' in nature.

Many of the October instrument questions are specifically included to provide information to score schools in terms of their SLM practices. Wills and van der Berg (2018) and Taylor, Wills and Hoadley (2018) provide a description of the instrument development and item writing process which was strongly underpinned by a leadership for literacy theoretical framework commissioned for this study (see Hoadley 2018).[4]

[4] Hoadley, A (2018) Leading for literacy: A review of the research. RESEP working paper. Stellenbosch University. http://resep.sun.ac.za/leading-for-literacy-a-review-of-the-research/

Taylor, N. and Hoadley, U. (2018) Leadership for Literacy: Exploring leadership practices in township and rural primary schools serving poor communities in South Africa. Final Report on the Case Study Schools. *RESEP Report.* http://resep.sun.ac.za/leadership-for-literacy-final-report-on-the-case-study-schools/

Wills, G. and van der Berg, S. (2018) Measuring leadership and management and their linkages with literacy in rural and township primary schools in South Africa. *Stellenbosch Working Paper Series No. WP21/2018. https://www.ekon.sun.ac.za/wpapers/2018/wp212018*

Table 10: Contextual and school leadership and management instruments administered - filenames for relevant datasets, instruments and codebooks

Instrument	Type of instrument	Grade applicability	When	Dataset filename	Instrument filename	Codebook filename
Grade 3 teacher interview	Face-to-face interview	Grade 3	Feb/Mar 2017 Oct/Nov 2017	gr3_interview_feb_clean.dta gr3_interview_oct_clean.dta	gr3teacher_feb_instrument.pdf gr3teacher_oct_instrument.pdf	codebook_gr3_teacher_feb.pdf codebook_gr3_teacher_oct.pdf
Grade 6 teacher interview	Face-to-face interview	Grade 6	Feb/Mar 2017 Oct/Nov 2017	gr6_interview_feb_clean.dta gr6_interview_oct_clean.dta	gr6teacher_feb_instrument.pdf gr6teacher_oct_instrument.pdf	codebook_gr6_teacher_feb.pdf codebook_gr6_teacher_oct.pdf
Educator survey	Self-completed survey	All educators in the school, all grades	Feb/Mar 2017 Oct/Nov 2017	educator_feb_clean.dta educator_oct_clean.dta	educator_feb_instrument.pdf educator_oct_instrument.pdf	codebook_ educator_feb.pdf codebook_ educator_oct.pdf
Principal interview	Face-to-face interview	n/a	Feb/Mar 2017 Oct/Nov 2017	principal_feb_clean.dta principal_oct_clean.dta	principal_feb_instrument.pdf principal_oct_instrument.pdf	codebook_principal_feb.pdf codebook_principal_oct.pdf
Deputy principal interview	Face-to-face interview	n/a	October 2017	deputy_oct_clean.dta	deputy_oct_instrument.pdf	codebook_deputy_oct.pdf
School Functionality	Observations: Self-completed by fieldworker.	n/a	Feb/Mar 2017 Oct/Nov 2017	func_feb_clean.dta func_oct_clean.dta	functionality_feb_instrument.pdf functionality_oct_instrument.pdf	codebook_func_feb.pdf codebook_func_oct.pdf
Teacher test	Self-completed test	Grade 6	February 2017	teachertest_feb_clean.dta	teachertest_feb_booklet.pdf	codebook_teachertest_feb.pdf
Classroom observation^	Observations: Self-completed by fieldworker.	Grade 3 & 6	Feb/Mar 2017 Oct/Nov 2017	class6obs_feb_clean.dta class6obs_oct_clean.dta	classobservation_feb_instrument.pdf classobservation_oct_instrument.pdf	codebook_classobs_feb.pdf codebook_classobs_oct.pdf
Book observation*	Observations: Self-completed by fieldworker.	Grade 3 & 6	Feb/Mar 2017 Oct/Nov 2017	bookobs_feb_clean.dta	booksobservation_feb_instrument.pdf booksobservation_oct_instrument.pdf	codebook_bookobs_feb.pdf

Notes:

*Data for the October book observations were only for a limited panel and have not been released with this data.

^The classroom observation instrument considers factors in the classroom that were observed outside of lesson time. This is not a lesson observation instrument.

7. General explanatory notes applying to all datasets

The next section is devoted to specific details of the individual datasets which users will find helpful when analysing the data.

7.1 Variable labelling conventions:

The following labelling conventions apply across all the datasets:

- Where variables are labelled with REFLECT EVIDENCE this indicates that the interviewer did not directly ask this question of the interviewee. Instead, the interviewer was asked to reflect on the respondent's answers to open-ended questions or multiple questions, or to reflect on documents provided by the school.
- Variables labelled with the preface "OBSERVATION" in the label are those questions that were based on observation of language books used to teach grade 3. These questions were not asked directly.
- Variables labelled with the preface "OFFICE ONLY" in the label (and in the question in the questionnaire) were completed by the researcher once the whole questionnaire had been reviewed.
- Variables labelled with the preface "DERIVED"in the label will not be found in the questionnaires as they have been derived from the data collected in the questionnaire.
- Variables labelled with the prefaced "Interviewer optional" provided a space for the interviewer to capture any further comments to clarify response on questions above. eg: Q4.7, Teacher interview, grade 6.

7.2 Actions taken to preserve anonymity

To preserve anonymity, the following variables were deleted from all data sets: School name, teacher name, date of interview, data collector name, date and time of interview.

Further actions were taken in individual data instruments, where required. These are detailed in the section of data notes on that instrument. Largely these took the form of excluding variables with open-ended string responses, where the responses could inadvertently identify the school, interviewer or respondent.

7.3 Linking of teacher data across datasets

A unique teacher id has been created to allow the linking of data by school and by teacher.

For Grade 6 teachers, the classroom observation (February & October), teacher test (February) and teacher interview (October) can be linked using schoolid and teacher_id. It is not possible to link this data to the Grade 6 teacher interview (February).

For Grade 3 teachers, the classroom observation (February & October) and teacher interview (October) can be linked using schoolid and teacher_id. It is not possible to link this data to the Grade 3 teacher interview (February).

The teacher_id is coded as:

g6_1 or g6_2, or g6_3 or g6_4 for grade 6 teachers and

g3_1 or g3_2 or g3_3 or g3_4, etc for grade 3 teachers.

The first part of the identifier indicates whether it applies to a grade 3 teacher (g3) or a grade 6 teacher (g6).

The second part of the identifier identifies the teacher. For example:

> A teacher_id of g6_1 was coded for the first Grade 6 teacher in a school whose classroom observed in February 2017.
>
> The teacher_id was coded as g6_2 for the second Grade 6 teacher in a school whose classroom was observed in February 2017 or where a second teacher did the teacher test.
>
> Where data across all instruments across the two time periods is for the same teacher, the teacher_id is g6_1 throughout.

In the next section additional explanatory notes are provided for each data instrument, in turn. Those instruments that were collected via Computer Assisted Interviewing are discussed first. For those instruments, information was directly captured onto tablets, using Tangerine software.

8. Additional explanatory notes for each dataset

8.1 Classroom observations – February

Instrument name: classroomobs_feb_instrument

Sample size:

Grade 6: 61
Grade 3: 82

Number of teachers per school:

Grade 6: One observation was done per school.
Grade 3: In 23 schools more than one observation per school

Method of collection:

Data directly captured onto tablet by interviewer, following observation of classroom.

8.1.1 Naming and coding conventions

- As this questionnaire was captured in Tangerine, there are no question numbers.
- "I don't know" is coded sequentially in some questions, but is coded as "9" in other questions. Eg: desk_size: Don't know is coded as 9; student_desk: Don't know is coded sequentially (coded as "3").
- *Shelving* (Is there sufficient shelving/storage in the classroom): The scale "Not enough at all" to "More than enough" is coded "1" to "3" with "4" for don't know, while for *student_desk* and *chairs* the coding for the same scale is as follows: "Not enough at all" to "More than enough" is coded "0" to "2" with "3" for don't know
- 'No' is coded as 0, Yes is coded as 1.

8.1.2 Further notes on specific questions

- o_grade (Grade for which classroom observation was done) shows whether the observation is for a grade 3 or 6 classroom.
- g6_tested (Is this the classroom of the grade 6 tested learners?) is coded 1 if the observation applies to the classroom of the grade 6 class which participated in the ORF tests. For all observations this is true
- g3_tested (Is this the classroom of the grade 3 tested learners?) is coded 1 if the observation applies to the classroom of the grade 3 class which participated in the ORF tests. For all observations this is true.
- Teacherid_g6 ("Unique teacher identifier: Grade 6"and Teacherid_g3 ("Unique teacher identifier: Grade 3") have been replaced by teacher_id, which identifies the teacher as a grade three or six teacher and further identifies the teacher uniquely within the school. Only one Gr6 teacher's classroom was observed per school, in February, thus all are coded as "1". This variable is important for linking this data with teacher data from other instruments. As more than one Grade 3 teacher's classroom was observation in some schools, this identifier is coded "1" for the first classroom observation for Grade 3., and "2" for the second classroom observation.
- A note on the variable *displayed*: This set of variables relates to displays in the Grade 3 classroom: Where classroom observations were done in grade 6 classrooms, displayed_1 to displayed_8 (which ask which of the following are displayed in Grade 3 classrooms) should be interpreted with caution.

8.1.3 Deleted variables

Class name was dropped to preserve anonymity.

8.2 Classroom observations – October

Instrument name: classroomobs_oct_instrument

Sample size:

52 for Grade 6.

77 for Grade 3.

Number of teachers per school:

One classroom observation per school for Grade 6.

For Grade 3: 22 schools where more than one Grade 3 classroom was observed.

Method of data collection:

Data directly captured onto tablet by interviewer, following observation of classroom.

8.2.1 Naming and coding conventions

As this questionnaire was captured in Tangerine, there are no question numbers.

- "I don't know" is coded sequentially in some questions and is coded as "9" in other questions. Eg: desk_size: Don't know is coded as 9; student_desk: Don't know is coded sequentially (coded as "3").

- *Shelving* (Is there sufficient shelving/storage in the classroom): The scale “Not enough at all” to “More than enough” is coded “1” to “3” with “4” for don’t know, while for *student_desk* and *chairs* the coding for the same scale is as follows: “Not enough at all” to “More than enough” is coded “0” to “2” with “3” for don’t know. Please note the difference in coding patterns.
- ‘No’ is coded as 0, Yes is coded as 1 (the same as February classroom observation)

8.2.2 Further notes on specific questions

- o_grade (Grade for which classroom observation was done) shows whether the observation is for a grade 3 or 6 classroom.
- g6_tested (Is this the classroom of the grade 6 tested learners?) is coded 1 if the observation applies to the classroom of the grade 6 class which participated in the ORF tests. For all observations this is true
- g3_tested (Is this the classroom of the grade 3 tested learners?) is coded 1 if the observation applies to the classroom of the grade 3 class which participated in the ORF tests. For all observations this is true.
- Teacherid_g6 (“Unique teacher identifier: Grade 6”and Teacherid_g3 (“Unique teacher identifier: Grade 3”) have been replaced by teacher_id, which identifies the teacher as a grade three or six teacher and further identifies the teacher uniquely within the school. This variable is important for merging classroom observation data with teacher test data (February) and teacher interview data (October)
- A note on the variable *displayed* : This variable relates to displays in the Grade 3 classroom. Where classroom observation were done in grade 6 classrooms, analysis of this variable is not sensible.

8.2.3 Deleted variables

Class name were dropped to preserve anonymity.

8.3 Book observations (February)

Instrument name: bookobservation_feb_instrument

Sample size: 398 learner books for learners in either Grade 3 or 6.

Method of data collection: Data directly captured onto tablet by interviewer, following observation of classroom.

8.3.1 Naming and coding conventions

There are no question numbers in this data as it was captured electronically using tangerine, which does not allocate question numbers.

- Pages_exercises (On how many pages in the CAPS/DBE English workbook 1 (for terms 1 & 2) has the learner completed any exercises: No CAPS/DBE book for learner is coded as 997. It is labelled as “No CAPS/DBE book for learner”.
- Days_exercises (For how many days since the beginning of the year is there evidence of any written work across all learner’s exercise books): No English exercise book(s) or learner is coded as 997. It is labelled as “Learner has no English exercise book”.
- Extended_writing: “Not enough time to count“ was coded as 995. It is labelled as “Not enough time to count”.

8.3.2 Further notes on specific questions

- Days_exercises (For how many days since the beginning of the year is there evidence of any written work across all learner's exercise books): The interviewer was instructed to count the days work was done in any of the English exercise books, but to exclude CAPS/DBE workbooks,
- Extended_writing: Interviewers were instructed to consider 5 or more sentences together as a paragraph as extended writing. Interviewers were instructed to count these examples of extended writing in both the English exercise book and English extended writing or free writing exercise book. Extended writing in the CAPS/DBE workbook were not counted.

8.3.3 Deleted variables

- Sch_name_other (Other school name) was dropped to preserve anonymity.
- Learner_name was dropped to preserve anonymity. Individual learners are identified by "Studentid".

8.3.4 Linking to other student data

Data from the book observation can be linked to data from the Oral Reading Fluency test (February), using schoolid and studentid.

8.4 Grade 3 teacher interview February

Instrument name: gr3teacher_feb_instrument

Sample size: n=61

Number of teachers instrument administered to per school: 1

Method of collection: Paper, transcribed into stata.

8.4.1 Naming and coding conventions:

- All variables in this file are prefaced with t3 indicating that this data was collected from the teacher in grade 3.
- 'Yes' is coded as "0". No is coded as "1"
- There are several placeholders in the data. For example: GR3_BACKGROUND_selfadminstered, GR3_READING_TEACHING. These separate the sections of the questionnaire, to make it easier to follow the order of the questionnaire and are empty variables.

8.4.2 Coding of "other"

Where there is the option to select "other", that variable is named numerically. For example Q1.3 "other" is named t3_subjects_7 as "other" is the 7th option for Q1.3.

The specifics of "other" have not been captured, in most cases. For example, teacher qualification (other) was not captured.

8.4.3 Further notes on specific questions:

- 0.6 The ESRC ID of the teacher was not captured. Unlike the October interview, there is no unique teacher identifier for this data.
- Q1.9 (Teacher's home language) – interviewers were instructed to choose at most 2 options.
- Q2.11 – In the last year was specified as Terms 2-4 of last year and 1st term of this year."

- Q6.6 A common test was defined as the same test for each subject that is given to all classes in the same grade.

8.4.4 Deleted variables

The following variables were deleted from the data sets.

Question no.	Question description	Reason for deletion	Data captured elsewhere?
0.4	Respondent cell	To preserve anonymity.	
Q1.8 Teacher qualifications	Other specify	No data captured	No.
Q1.8 Teacher qualifications	Any further comments to clarify response on Q1.1 to Q1.8?	To preserve anonymity.	
Q4.7	Further comments on Q4	To avoid duplication of data	Q4.1 to 4.5
Q5.3	Further comments on Q5.1 & Q5.2	To avoid duplication of data	Q5.1 & Q5.2
Q6.2	Assessments: Can you explain what norms or benchmarks or criteria are?	To preserve anonymity.	
Q6.5	Any further comments to clarify responses on assessment plans?	To preserve anonymity.	
Q6.8	Any further comments to clarify responses on internal common tests?	To preserve anonymity.	

8.4.5 Linking this data to other teacher datasets for this project:

This data cannot be linked to other teacher data at this stage.

8.5 Grade 3 teacher interview October

Instrument name: gr3teacher_oct_instrument

Sample size: n= 63

Number of teachers per school: In 3 schools, more than one educator was interviewed.

Method of collection: Paper, transcribed into stata.

This questionnaire differs from the questionnaire administered in February 2017. In most cases, the same teachers were interviewed in February and October.

8.5.1 Naming & coding conventions:

- All variables in this file are prefaced with t3 indicating that this data was collected from the teacher in grade 3.
- 'No' was coded as "0", Yes was coded as "1".
- Where the same question is repeated, for different respondents, variable names are repeated, but the number "1" is appended to the variable name, for the second variable.

Example: t3_qual2 is the variable describing the teacher's qualification for the home language teacher (Q1.5) whereas t3_qual21 is the variable describing the teacher's qualification for the FAL teacher (Q2.3).

8.5.2 Coding of "other"

- Where an "other, specify" option was allowed in a single mention question, and where "specify" was captured, the main variable is numbered "2" and responses under the relevant "other, specify" is numbered "1".

  Example: t3_qual2 captures the teachers qualification (one of the pre-coded options) and t3_qual1 captures "other, specify" for this question.

8.6 Further notes on specific questions:

- Q2.1 to Q2.4 is the same as Q1.1 to Q1.4, but Q2.1 to Q2.4 applies to the First additional language teacher, while Q1.1 to Q1.4 applies to the Home language teacher.
- Q21.9 is coded as a multiple mention question, although there are no specific instructions to this end in the questionnaire.

8.6.1 Deleted variables

The following variables were deleted from the data sets.

Question no.	**Question description**	**Reason for deletion**	**Data captured elsewhere?**
Q4.5		no observations	
Q8.1.b	How many learners have access to DBE workbooks in African language?	No data captured	
Q8.2.a	How many learners have access to textbooks in African language	No data captured	
Q9.1 c.	How many learners have access to textbooks in English?	No observations	
Q9.2	Which do you allow learners to take home)?: option 1 (textbooks)	No observations	
Q10.1	What type of books mostly used for teaching English)?: option 0 (textbooks)	No observations	
Q11.1	other(specify) options	To preserve anonymity	
Q11.3	Which teacher is best at teaching?	To preserve anonymity	

The questions in the next table were not directly captured and have been deleted. However, the data from these questions is captured elsewhere as shown. Many of the questions below were asked and recorded as open-ended questions, but responses were captured by the interviewer (in the case of a REFLECT question) and by the researcher (in the case of an OFFICE ONLY question). :

Question no.	Question description	Reason for deletion	Data captured elsewhere?
Q10.2	Anyone in the school management team responsible to reading?	Not captured directly	Q10.3 (t3_smtread2),
Q12.1	(Favourite children's book)	Not directly captured.	Q21.3 (An OFFICE ONLY question)
Q14.4	Open-ended question on possible reasons for some learners struggling to read	Not directly captured.	Q21.7
Q14.5		Data in Q21.7 (OFFICE ONLY) is more reliable	Q21.7
Q15.3	Open-ended question on what teacher learned at reading training	Not directly captured.	Q15.4 (REFLECT)
Q16.1		Not directly captured.	Q16.3 (REFLECT)
Q16.2		Not directly captured.	Q16.3 (REFLECT)
Q17.1		Not directly captured.	Q17.2 & Q17.3 (REFLECT)

8.7 Grade 6 teacher interview (February)

Instrument name: gr6teacher_feb_instrument

Sample Size n = 61 schools.

Number of teachers instrument administered to per school: 1

Method of collection: Paper, transcribed into stata.

To save time in administering the questionnaire, teachers were allowed to self-complete Q1 – Q3 of the questionnaire (background, reading and teaching resources, own reading.

This questionnaire was missing for 3 of the 61 schools.

8.7.1 Naming and coding conventions

- All variables in this file are prefaced with t6 indicating that this data was collected from the teacher in grade 6.
- 'Yes' is coded as "0". No is coded as "1". This differs from the February Grade 3 teacher interview
- Where an option was provided for "other", this is generally named with the suffix _other. For example: Q1.3 (subjects taught) is named t6_subjects_other.

8.7.2 Coding of "other"

- Q1.7 Teacher qualifications. Variable t6_qualification_other captures whether the respondent selected the "other qualification" option. However the qualification specified was not captured.

- Q1.9 Teacher's home language. The data specifying the “other” language was not captured. t6_lang_other merely captures whether the teacher’s home language is anything other than the language listed.

8.7.3 Further notes on specific variables

- Q1.9 – Teachers were instructed to choose at most two languages as home language.
- Q5.7 The question included the instruction: Very poor learners would be the poorest in THIS school.
- Q5.10: The first option (Not much. There are too many students in my class to support individual students) was not captured.
- Q6.6 common tests was specified as: the same test for each subject is given to all classes in the same grade.

8.7.4 Deleted variables

The following variables were deleted from the dataset for the reasons shown:

Question no.	**Question description**	**Reason for deletion**	**Data captured elsewhere?**
Q1.8	Further comments on teacher background	To preserve anonymity	
Q4.7	Further comments on differentiation in the classroom	To preserve anonymity	
Q5.3	Further comments on time spent on reading activities	To preserve anonymity	
Q5.9	Further comments to clarify assistance to learners who are struggling to read	To preserve anonymity	
Q6.4 ,		To preserve anonymity	
Q6.4 ,		To preserve anonymity	
Q6.10		To preserve anonymity	

8.7.5 Linking this data to other teacher datasets

ESRC ID of teacher was not captured. This data cannot be linked to other teacher data at this stage.

8.8 Grade 6 teacher interview (October)

Instrument name: gr6teacher_oct_instrument

Sample size: n= 60

Number of teachers per school: One educator per school was interviewed.

Method of collection: Paper, transcribed into stata.

This questionnaire differs from the questionnaire administered in February 2017. In most cases, the same teachers were interviewed in February and October.

8.8.1 Naming and coding conventions:

- All variables in this file are prefaced with t6 indicating that this data was collected from the teacher in grade 6.
- 'No' is coded as "0".Yes is coded as "1". This differs from the February Grade 6 teacher interview.
- Where there is an "other, specify option" in a multiple mention question, these responses are generally found in a variable with the extension " _other".

8.8.2 Further notes on specific questions

- Q1.5 (Teacher qualifications) was asked as a multiple mention (with the instruction to tick all), but has been coded in t6_qual1 as a single mention. Where the respondent chose "other, specify", the responses are coded in t6_qual2.
- Q15.6 is coded as a multiple mention question, although there are no specific instructions to this end in the questionnaire.
- Q15.7 – Interviewers were given the following instructions: "Very vague" would be for example just mentioning reading or writing or that more children can read well or write well. Clearer responses would describe what children can read, or the marks obtained in a specific test, or words read correctly per minute, or expression in their voice etc. They may mention the writing of paragraphs, extended texts).

8.8.3 Deleted variables

The following variables were deleted from the data sets for the reasons shown:

Question no.	Question description	Reason for deletion	Data captured elsewhere?
Q2.5	Interviewer optional comment	Not captured	No
Q5.6	Interviewer optional comment	Not captured	No
Q4.3, option 0 (textbooks)		No observations	NO
Q4.4 option 1 (textbooks)		No observations	No

The questions in the table on the next page were not directly captured and have been deleted. However, the data from these questions is captured elsewhere as shown. Many of the questions below were asked and recorded as open-ended questions, but responses were captured by the interviewer (in the case of a REFLECT question) and by the researcher (in the case of an OFFICE ONLY question).

Question no.	Question description	Reason for deletion	Data captured elsewhere?
Q5.1	(Which teacher) is best at teaching reading	to preserve anonymity.	Q5.5 (REFLECT)
Q5.3	Person teachers could go to for expert support in teaching reading to learners who are struggling to read.	to preserve anonymity.	Q5.5 (REFLECT)
Q6.1	Title of the respondent's favourite children's book	Not directly captured	Q15.2 (OFFICE ONLY)
Q6.2	Last novel teacher read	Not directly captured	Q6.3 (REFLECT)
Q8.2	How do you go about Group Guided Reading?	Not directly captured	Q8.4 and Q15.4 (OFFICE ONLY) Respondent's understanding of Group Guided Reading.
Q8.3	Why do you think (Group Guided Reading) is or isn't a useful method of teaching reading?	Not directly captured	
Q9.4	Possible reasons for why some learners struggling to learn to read	Not directly captured	Q9.5 (REFLECT) & Q15.5 (OFFICE ONLY)
Q10.3	What have you implemented from training	Not directly captured	Q10.4 (REFLECT)
Q11.1	Open-ended: improving reading outcomes	Not directly captured	Q11.3 (REFLECT)
Q11.2	Open-ended	Not directly captured	Q11.3 (REFLECT)
Q12.1		Not directly captured	Q12.2 & Q12.3

8.9 Educator survey interview (February)

Instrument name: educator_feb_instrument

Sample size: 967 educators

Number of teachers per school: Multiple educators per school.

Method of collection: Paper, transcribed into stata.

This questionnaire was anonymously self-completed by all educators in the school.

8.9.1 Naming and coding conventions

- All variables in this file are prefaced with t_ indicating that this data was collected from a teacher.
- BACKGROUND, LEADERSHIP_VISION, READING_TEACHING, ABSENTEEISM_TIME, SUPPORT_ACCOUNTABILITY, CONSULTATION_COMM are placeholders in the data.
- YES is coded “0”; NO is code as “1” in this dataset.

8.9.2 Further notes on specific questions

- Q3.1. Although the instruction was to choose at least 2, only one response was provided and the question was coded as such.
- Q5.4 Principal and Q5.5 (HOD) observe teaching in classroom. The question asked specifically for cases where the principal (or HOD) observes teaching for at least 10 minutes of a lesson.
- Q8.5 The instruction was given: including union meetings or any union supported gatherings, including memorial services.

8.10 Educator survey interview (October)

Instrument name: educator_oct_instrument

Sample size: 957 educators

Number of teachers per school: Multiple educators per school.

Method of collection: Paper, transcribed into stata.

This questionnaire was anonymously self-completed by all educators in the school (n=957 educators).

8.10.1 Naming and coding conventions

- All variables in this file are prefaced with t_ indicating that this data was collected from a teacher.
- No details written under “specify” were captured for any other (specify) options in this questionnaire. Instead the data records simply whether the responded selected “other”.
- NO is coded “0”; YES is code as “1” in this dataset. This differs from February.

8.10.2 Further notes on specific questions

- Q3.1 (Reason why NEW textbooks or readers chosen). No data was collected for the options 3.1_5 “The content is matched to CAPS” and option 3.1_6 “The content is matched to the ability of learners” and these options have been deleted from the dataset.

- Q4.4 and Q9.3 pertains to lesson observation for at least 10 minutes (for reasons other than IQMS). A specific instruction was included in both questions to this regard.

8.11 Principal interview (February)

Instrument name: principal_feb_instrument

Sample size n = 61 schools.

Number of teachers per school: 1 principal per school.

Method of collection: Paper, transcribed into stata.

There were high levels of missing data for some schools for this questionnaire.

In 4 schools there was principal and other staff members were interviewed.

8.11.1 Naming & coding conventions

- All variables are prefaced with p_
- All other(specify) are captured as p_varname_other. For example: p_permanent_other captures the specified answers where “other” was selected at Q1.1.
- In this dataset "Yes" is coded as 0. "No" is coded as 1.

8.11.2 Further notes on specific questions

- Q2.3 (How often are their conflicts between teachers & school management?) There are two codes for “no response” (5 & 6).
- Q11 is asked of the principal. If the principal is unable to answer, the same questions were asked of the deputy principal in Q12.
- Q13 could be asked of either the principal or deputy.
- Q14.3 was asked as a multiple mention (Interviewer was instructed to tick all), but single responses were given and the question was coded as a single mention question.
- Q18.4 (How regularly are teacher rewards/awards given?): Option 1 (Adhoc) included the examples: once off event, when the principal feels like it, when someone donates for the awards.
- Q18.5 (How do you decide which teachers get an award?) – N/A is coded as “997”
- Q22 – Q23 could be asked of an administrator, deputy or principal. For this reason Q21 repeats school identifying information and was dropped to preserve anonymity.

8.11.3 Deleted variables

The following variables were deleted from the data sets for the reasons shown:

Question no.	Question description	Reason for deletion	Data captured elsewhere?
Q5.2	Other (specify)	Only 1 response given	No
0.4	Principal cellphone	To preserve anonymity	
0.5	Principal email	To preserve anonymity	
Q6.4	If you are not interviewing the principal explain why?)	To preserve anonymity	
Q8.5		To preserve anonymity	
Q9.9	Any further comments on principal's clarity on which learners and classes are weakest	To preserve anonymity	
Q10.10	Any further comments on SMT meeting minutes	To preserve anonymity	
Q15.4	Further comments on Q15	To preserve anonymity	
Q16.3	Further comments on attracting good teachers to the school, and principal's level of influence	To preserve anonymity	
Q17.6	Further comments on identifying best teachers	To preserve anonymity	
Q18.1	Monetary or non-monetary rewards/awards. Other (specify)	To preserve anonymity	
Q18.7	Further comments on awarding best teachers.	To preserve anonymity	
Q19.2	Tell me about the last time you had to address problems of poor teaching in this school	To preserve anonymity	
Q19.8	Further comments	To preserve anonymity	
Q20.1	(Greatest challenges you face as principal)	To preserve anonymity	

8.12 Principal interview (October)

Instrument name: principal_oct_instrument

Sample size: 60

Number of teachers per school: 1 principal per school.

Method of collection: Paper, transcribed into stata.

8.12.1 Naming & coding conventions

- All variables are prefaced with p_
- "No" is coded as 0, "Yes" is coded as 1
- Q7.1, Q8.1 and Q8.2, Q9.1 and Q9.2: are prefaced REFLECT EVIDENCE or ASK. Where documents were available, the answers to these questions would come directly from the financial documents. Interviewer were instructed that if the school could not provide the relevant financial documents (income and expenditure statement or budget), the interviewer asked for an estimate of these amounts.

8.12.2 Further notes on specific questions

- Data for Q4.5 should be considered together with Q11.1 (po_smtvacant4). Data in Q11.1 should be considered more reliable than in Q4.5.
- Q5.6 should be considered together with Q11.2 (po_ownbooks2). Data in Q11.2 should be considered more reliable than in Q5.6.

8.12.3 Deleted variables

The following variables were deleted from the data sets for the reasons shown:

Question no.	Question description	Reason for deletion	Data captured elsewhere?
Q2.2	How often does this reading period actually happen?	Not captured	No
Q6.2	Interviewer optional: Any comments on Financial documents shown?)	To preserve anonymity.	No
Q9.5	Interviewer: Any useful comments to help clarify Q9?)	Not captured	No

The following questions were not directly captured and have been deleted. However, the data from these questions is captured elsewhere as shown. Many of the questions below were asked and recorded as open-ended questions, but responses were captured by the interviewer (in the case of a REFLECT question) and by the researcher (in the case of an OFFICE ONLY question).

Question no.	Question description	Reason for deletion	Data captured elsewhere?
Q4.4	If there are vacancies is there anyone who is doing the job?	Not directly captured	Q4.5 (REFLECT) Q11.1 (OFFICE ONLY)
Q5.5	Have you ever tried to apply to get certain functions to pay for your own books?	Not directly captured	Q5.6 (REFLECT) Q11.2 (OFFICE ONLY)

8.13 Deputy Principal Interview: October

Instrument name: deputy_oct_instrument

Sample size: 60

Number of teachers per school: 1 deputy-principal per school.

Method of collection: Paper, transcribed into stata.

This interview was only done in the October round of fieldwork.

8.13.1 Naming & coding conventions

- All variables are prefaced with dp_
- In this dataset, Yes is coded as '1', No is coded as '0'.

8.13.2 Further notes on specific questions

- Q1.4 is coded as a multiple mention, although the question is worded as a single mention. Several schools mentioned more than one LOLT in the Foundation Phase.
- Q2.3 should be considered together with Q11.1. The data in Q11.1 should be considered more complete.
- *Q2.3 and Q2.6: Note to interviewer: Very vague would be for example just mentioning reading or writing or that more children can read well or write well. Clearer responses would describe what children can read, or the marks obtained in a specific test, or words read correct per minute, or expression in their voice etc. They would mention the writing of paragraphs, extended texts."*
- Q4.2 and Q4.4: Although the question was asked as a multiple mention (interviewers were instructed to tick all that apply), only one response was provided per respondent and the data was coded as a single mention.
- Q2.6 should be considered together with Q11.2. The data in Q11.2 should be considered more complete.
- Q3.6 (How much has the principal tried to get monies or resources donated to this school?) should be considered together with Q11.3. The data in Q11.3 should be considered more complete.
- Q6.5 should be considered together with Q11.4 (OFFICE ONLY Which of the following best describes the response to the questions on identifying reading specialist?). The data in Q11.4 should be considered more complete.
- Q8.7 was asked as a multiple mention (interviewers were instructed to tick all that apply), but only one response was provided per respondent and the data was coded as a single mention.
- Data from Q11.7 should be considered more complete and reliable as it takes the whole questionnaire into consideration and was completed by the researcher.

8.13.3 Deleted variables

The following variables were deleted from the data sets for the reasons shown:

Question no.	Question description	Reason for deletion	Data captured elsewhere?
Q1.2	For how long has deputy principal been working in current position?	Not captured	No
Q1.3	For how long has deputy principal been working in current school?	Not captured	No
Q4.5	Further comments on fundraising	To preserve anonymity.	
Q6.6	Interviewer optional: Further comments on reading specialists	Not captured	

The following questions were not directly captured and have been deleted. However, the data from these questions is captured elsewhere as shown. Many of the questions below were asked and recorded as open-ended questions, but responses were captured by the interviewer (in the case of a REFLECT question) and by the researcher (in the case of an OFFICE ONLY question).

Question no.	Question description	Reason for deletion	Data captured elsewhere?
Q2.1	Tell me about any learning or language goals set for Grade 3.	Not directly captured	Q2.2 (REFLECT) Q2.3 (REFLECT) Q11.1 (OFFICE ONLY)
Q2.4	Tell me about any learning or language goals set for Grade 6	Not directly captured	Q2.5 (REFLECT) Q2.6 (REFLECT) Q11.2 (OFFICE ONLY)
Q3.1	Tell me about a time the principal tried to get monies or resources donated to this school	Not directly captured	Q3.7 (REFLECT) Q3.8 (REFLECT)
Q3.3	What was involved in fundraising events in the last 2 years?	Not directly captured	Q3.7 (REFLECT) Q3.8 (REFLECT)
Q3.4	Roughly how much was donated	Not directly captured	Q3.8 (REFLECT)
Q5.5	Did (deputy principal) ever specialise or receive training in teaching of reading or language?	Not captured	No
Q5.6 (REFLECT)	Which of the following best describes whether the respondent actually specialised or received training in teaching reading or language	Not captured	

Q6.1	Who would you say is best at teaching reading in this school or at each grade?	Not captured as they contained teacher names	Q6.5 REFLECT Q11.4 (OFFICE ONLY)
Q6.3		Not captured as they contained teacher names	Q6.5 REFLECT Q11.4 (OFFICE ONLY)
Q7.1	Which of the following best describes how the roles and responsibilities of deputies are defined	Not directly captured	Q7.3 (REFLECT)
Q7.2	Which of the following best describes how the roles and responsibilities of HODs are defined	Not directly captured	Q7.4 (REFLECT).
Q7.5	Is anyone in the management team responsible for overseeing reading	To preserve anonymity	Q7.6 and Q7.7 (REFLECT).
Q8.3	An open-ended question: Are there currently any programmes to support reading in this school?")	Not captured	Q8.7 (REFLECT)
Q9.1	What do you think is the ONE MAIN thing that needs to be done to improve reading outcomes in the school	Not directly captured	Q9.3 (REFLECT How much understanding does the respondent seem to have about how to improve reading outcomes?) Q11.7 (OFFICE ONLY).
Q9.2	Imagine a school where most learners can't read and teachers do not know how to teach reading	Not directly captured	Q9.3 (REFLECT How much understanding does the respondent seem to have about how to improve reading outcomes?) Q11.7 (OFFICE ONLY).

8.14 School Functionality: February

Instrument name: functionality_feb_instrument

Sample size: 61

Number per school: 1 functionality assessment per school.

Method of collection: Paper, transcribed into stata. No direct interviewing. All data is based on the interviewer's observations

8.14.1 Naming & coding conventions

- All variables are prefaced with f_
- Throughout the questionnaire, not applicable is coded as 997. For example: Q2.4b. Where not applicable, this is coded as 997.
- Yes is coded as 0, No is coded as 1.

8.14.2 Further notes on specific questions

- Q2.1 Number of classrooms. Definition of classroom used: Any room that can reasonably accommodate students in class and a teacher. Mobile classrooms and containers converted into classrooms should also be counted as a classroom.
- Q4.1 "specify" was not captured for those who selected other, specify at Q4.1.
- While the data for Q3.2 to Q3.6 has been left in the dataset, users are encouraged to use the final data on classroom counts in Q3.7 to Q3.11 as it is more reliable. The same applies to Q 5: Data in Q5.7 to 5.11 is more reliable than data in Q5.2 to 5.6.
- Q3.9 & Q5.8 Interviewer instruction was: The teacher is considered not-teaching if they are not doing any productive educational activity e.g. eating, just sitting, marking, chatting to a friend, etc.
- Q3.11 & Q5.11: Interviewers were instructed that if there are one or two teachers in the classroom but no students, these classrooms were to be counted as not utilised.
- Q4.1 Fieldworker instruction: School-feeding related activities include for example any teachers or students leaving classrooms to collect utensils, food or water. The grade 6 fieldworker was asked to keep a track of this.
- Q6.2 – If staff toilets are not separate from learner toilets this was coded as "N/A. None"
- Q6.11 to 6.12 apply to school library, not classroom corner or mobile libraries.

8.14.3 Deleted variables

The following variables were deleted from the data sets for the reasons shown:

Question no.	Question description	Reason for deletion	Data captured elsewhere?
Q3.13		To preserve anonymity.	No
Q5.13		To preserve anonymity.	No
Q6.14		To preserve anonymity of field workers.	No
Q7.15		To preserve anonymity.	No. Information related to the data collection process, not to the actual data collected

.

8.15 School Functionality: October

Instrument name: functionality_oct_instrument

Sample size: 60

Number per school: 1 functionality assessment per school.

Method of collection: Paper, transcribed into stata. No direct interviewing. All data is based on the interviewer's observations

8.15.1 Naming & coding conventions

All variables are prefaced with f_

No is coded as 0, Yes is coded as 1 – this differs from the February school functionality dataset.

8.15.2 Further notes on specific questions

- Q1.2 The interviewer was provided with the following definition of a classroom: Any room that can reasonably accommodate students in class and a teacher. Mobile classrooms and containers converted into classrooms should also be counted as a classroom.
- Q2.1. "specify" was not captured for those who selected other, specify at Q2.1.
- Q1.3 School feeding related activities include for example any teachers or students leaving classrooms to collect utensils, food or water.
- Q3.7: teaching includes all productive educational activity e.g. actively invigilating examination, actively supervising learners' reading, talking with learners. It excludes marking at desk.
- Q3.8 Teacher is considered not-teaching if not doing any productive educational activity e.g. eating, just sitting, marking, chatting to a friend, etc.
- Q3.11 E. If there were one or two teachers in the classroom, but no students, this was not counted as being utilised.
- Q4.1 Posters were described as including visual displays of reading events or notices about competitions. The fieldworker was instructed to look in corridors, open areas, reception, library, rather than classrooms.
- Q4.3 A further instruction was given to fieldworkers – by up-to-date meaning do the posters refer to recent events or upcoming events or would be useful for encouraging reading at the present time?
- Q5.3 In determining the length of the school-wide reading period: If different times were allocated for reading each day, the interviewer was instructed to choose the most common duration.

8.15.3 Deleted variables

The following variables were deleted from the data sets for the reasons shown:

Question no.	Question description	Reason for deletion	Data captured elsewhere?
Q2.4	Further comments on breaks/lunchtime and school feeding	no observations	No
Q3.1	Time interviewer started walk around	Not captured	No
Q3.13		No observations	No

Q3.2 to Q3.6	These were intermediate questions – a space for the interviewer to enter the tally of classrooms where teachers were present and teaching, present, but not teaching, absent, etc	Not captured	Q3.7 to 3.11
Q8.2	What time did the majority of teachers leave	No observations	
Q8.4	Reasons school teaching disrupted by some event	Not captured	
Q8.11	Further comments	Not captured	

8.16 Teacher test

Instrument name: teachertest_feb

Sample size: 61

Number per school: 1 teacher test per school.

Method of collection: Teachers provided answers to the test on paper. These were marked and the marks were imported into stata.

This datafile contains teachers' scores on the vocabulary test that was administered to Grade 6 teachers. The data does not detail the answers given by teachers, but rather whether answers provided were correct. The same test was administered to Grade 6 learners and to their teachers.

This test was administered once, in February 2017.

8.16.1 Naming & coding conventions

Most variables in this datafile are prefaced with tt_ , which stands for teacher test.
For each question, answers were scored as 1, 0.5 or 0. A correct answer was coded as 1, an incorrect answer was coded as 0 and a partially correct answer was scored as 0.5.

8.16.2 Further notes on specific questions

- tt_total is the teacher's overall total score. The total possible score on the test is 72.
- The teacher test was organised in sections: Section A to D. Questions were numbered sequentially within these sections.
- tt_total_a is a sub-total: The teacher's score for Section A.
- tt_total_b is a sub-total: The teacher's score for Section B.
- tt_total_c is a sub-total: The teacher's score for Section C.
- tt_total_d is a sub-total: The teacher's score for Section D.
- The total possible score for each section is 18.
- a_1 to a_18 are the scores for each question in section A, (Question 1 to 18).
- b_1 to b_18 are the scores for each question in section B, (Question 1 to 18).
- c_1 to c_18 are the scores for each question in section C, (Question 1 to 18).
- d_1 to d_18 are the scores for each question in section D, (Question 1 to 18).