RESEP
Research on Socio-Economic Policy

Stellenbosch
UNIVERSITY
IYUNIVESITHI
UNIVERSITEIT

forward together
sonke siya phambili
saam vorentoe

WHAT RICH NEW EDUCATION DATA CAN TELL US

New insights into learner flows, assessment, learner mobility and the subjects teachers teach

Authors

Servaas van der Berg, Chris van Wyk, Martin Gustafsson, Herman Meyer, Abigail Chari, Chloé van Biljon, Adaiah Lilenstein, Rebecca Selkirk & Jessica McCallum

Michael & Susan Dell FOUNDATION

We gratefully acknowledge funding support from the Michael & Susan Dell Foundation

CONTENTS

Authors:
Servaas van der Berg, Chris van Wyk, Martin Gustafsson, Herman Meyer, Abigail Chari, Chloé van Biljon, Adaiah Lilenstein, Rebecca Selkirk & Jessica McCallum

Design and layout:
Ink Design Publishing Solutions, Cape Town, www.inkdesign.co.za

Cover photo:
Ismail Kohail/Vecteezy.com

December 2023

We gratefully acknowledge support from the Michael & Susan Dell Foundation

Michael & Susan Dell FOUNDATION

List of Figures

List of Tables

What rich new education data can tell us: New insights into learner flows, assessment, learner mobility and the subjects teachers teach

EXECUTIVE SUMMARY

Overview

This report leverages data mainly derived from the South African School Administration Management System (SA-SAMS), including a unique longitudinal version of the Data Driven Districts (DDD) data for three provinces as well as the Learner Unit Record Information Tracking System (Lurits) data and National Senior Certificate (matric) examination data to analyse the dynamics of the South African education system in the aftermath of the Covid-19 pandemic. It provides a detailed analysis of learner flows, assessment strategies, and teacher dynamics, offering valuable insights for policymakers and educational stakeholders and illustrating how such data can be used in education policy and planning.

Key Findings

1. **The education data system (Chapter 1)**
   - The SA-SAMS and LURITS data systems have revolutionized the tracking and analysis of learner and teacher data, enabling a nuanced understanding of educational trends at various levels across South Africa. In addition, data from the matric examinations over more than a decade allows an analysis of important trends.

2. **Assessment changes and learner flows (Chapter 2)**
   - **Assessment leniency:** To mitigate pandemic disruptions, assessments were relaxed, resulting in decreased repetition rates, particularly in grades 10 and 11.
   - **Improved school-based assessment marks and greater learner progression:** This leniency improved SBA marks and led to more learners advancing to higher grades and a significant reduction in school dropout rates.
   - **Performance in key subjects:** A detailed analysis was conducted on the relationship between school-based assessments and NSC examination results in key subjects.
   - **Rise of high-performing poor schools:** The report investigates the phenomenon of poorer schools (Quintiles 1–3) excelling beyond expectations in matric results.

3 **Learner mobility analysis (Chapter 3)**

- **High mobility in grade 10:** Grade 10 witnessed the highest rates of inter-school mobility, indicating significant transitions to the Further Education and Training (FET) Phase.
- **Stability in grade 11:** In contrast, grade 11 experienced the lowest mobility, suggesting greater stability at this level.
- **Mobility among repeaters:** Repeaters exhibited slightly higher mobility than non-repeaters, reflecting the challenges faced by students who need to repeat a grade.
- **Consistent inter-provincial mobility:** Despite the pandemic, inter-provincial mobility remained stable, with minimal impact from the pandemic's disruptions.
- **Shifts in public and private school enrolment:** The pandemic reversed the flows of learners from public to private schools, except in Gauteng and the Western Cape, supporting anecdotal evidence that financial distress influenced school choice during the pandemic.
- **Visual mobility trends:** The use of visual tools like dot density and heat maps enhance the understanding of these mobility trends.

---

4 **New analysis of teacher data (Chapter 4)**

- **SA-SAMS as a National Educator Dataset:** SA-SAMS has been instrumental in building a comprehensive national dataset of individual educators.
- **Educator specialisation:** The report provides insights into educator specialisation in various subjects, which has been a largely unexplored area.
- **Data completeness and reliability:** Examines the completeness and reliability of the educator data, with a high degree of internal consistency noted.
- **Challenges in data usage:** Highlights the limitations in using the data for calculating attrition rates due to underrepresentation of certain educator demographics.
- **Statistics by specialisation:** Presents detailed statistics on educator specializations across phases and subjects, enhancing understanding of teaching patterns.
- **Future directions:** Concludes with reflections on further potential uses of SA-SAMS educator data for system-wide monitoring and planning.

---

5 **Conclusion and looking ahead (Chapter 5)**

- This report provides a snapshot of the current state of education in South Africa and lays the groundwork for future inquiries and strategic interventions.
- Leveraging data-driven insights and fostering collaborative data-gathering and analysis efforts among stakeholders in education are essential elements in building an education system that is robust, equitable and responsive to the needs of all its participants.

This report underscores the importance of the vital data-systems that are starting to provide major new insights into the functioning and growth of the education system and the complex effects of the pandemic on it, ranging from learner progression to teacher dynamics. Comprehensive data analysis serves as a powerful tool in understanding these effects and guiding future education policies and strategies. The insights garnered are vital for teachers, policymakers and stakeholders in navigating the post-pandemic educational landscape in South Africa.

CHAPTER 1
INTRODUCTION

> This report aims to demonstrate the potential of these and other datasets.

1.1 Background

The South African education data system has undergone significant transformation since the political transition. Collecting of individual data within the South African School Administration Management System (SA-SAMS) has been central in this regard. The current data infrastructure now facilitates comprehensive analyses crucial for understanding educational dynamics and shaping policy. SA-SAMS contains data on all learners and teachers and is administered at school level before being collated at district and provincial level for provincial planning purposes. It is also further used to create the Lurits data, a national data set designed for system coherence that also offers possibilities of tracking learners over time. The SA-SAMS data is used in the Data Driven Districts programme, an initiative by the Department of Basic Education in partnership with the Michael & Susan Dell Foundation and their technical partners. On request of Resep, this data has also now been linked across years for three provinces and anonymised before being provided to Resep for analysis. This analysis of learner performance and trends therefore results from cooperative efforts within the South African education community at different levels, from the school level to districts, provinces and the national Department of Basic Education (DBE). The insights drawn from this are invaluable for teachers, officials and policymakers in identifying areas of need and crafting targeted interventions to support learners or schools that may be struggling.

This report aims to demonstrate the potential of these and other datasets. The longitudinal data from Gauteng, the Eastern Cape and Limpopo enriches our understanding of various facets of the education system. In addition, data was available from the National Senior Certificate or matric exams, as well as from Lurits, the national data set based on the SA-SAMS. SA-SAMS data from KwaZulu-Natal is also included in the analysis of teacher data in Chapter 4. Teacher data has always been collected by schools within the SA-SAMS system, but has never been used for administrative or planning purposes. Thus it has not yet experienced the quality improvement

that often follows extensive use and interrogation of the data. Nevertheless, this data for the first time allows an analysis of teachers and the subjects they teach. This is information that provincial planners have been clamouring for. It also potentially allows analysis of the subject specialisation of older teachers that are likely to form part of the retirement wave now affecting the South African education system.

This research project follows previous projects, also funded by the Michael & Susan Dell Foundation (MSDF), that aimed to use high-quality administrative data to understand the performance of South Africa's schooling system.[1] More and improved data makes it possible to extend previous analyses regarding some important features of the South African education system:

DATA AVAILABILITY AND THE INCREASED GENERATION AND LINKING OF EDUCATION DATA

The analysis in this report is based on four datasets, namely the Learner Unit Record and Tracking System (Lurits) data, data from the Data Driven Districts (DDD) initiative, and National Senior Certificate (NSC) examination data. The DDD data is SA-SAMS data regularly collected from schools as part of the Data Driven Districts programme, an initiative by the Department of Basic Education in partnership with the Michael & Susan Dell Foundation and their technical partners. This data is then used to populate Data Driven Districts dashboards, a potential management tool. For this report, DDD data from different years were matched and linked over time to create a longitudinal dataset, and then anonymised and made available to Resep for this analysis. The DDD data available for this research covered only the Eastern Cape, Limpopo and Gauteng. These three provinces nevertheless, according to the Masterlist of Schools for 2021, represent 45 percent of all enrolments and 48 percent of the public schools in the country (DBE, 2022). While these provinces are not fully representative of the situation nationally, they give a good picture of what is likely happening at the national level.

While the quality of the data stored in SA-SAMS has improved considerably in recent years, there are still various data quality issues that could be improved. Firstly, the available SA-SAMS data obtained for the DDD still has some inconsistencies, as not all schools submit their data every term. Since cohort analysis requires data from the same schools over time, in some instances only data from schools that submitted every year can be used to conduct such analysis. Although many schools in Gauteng submitted learner-level data each year for the period 2016–2023, the number that submitted varied greatly between years, with low levels of submission in the earlier years. Analysis of longitudinal data presented in this report is, unless stated otherwise, only based on the sub-sample of schools that submitted data each year in the three provinces considered.

The quality and comprehensiveness of the LURITS and DDD data have improved over the period 2019 to 2022. For instance, the share of duplicate learner records collected decreased substantially between 2019 and 2022. This indicates that information was captured with increased accuracy, resulting in improved data quality which can also enhance the accuracy of analysis.

1 Van der Berg et al., 2019, 2020, 202s, 2022.

Datasets	Purpose	Comment
DDD	Enrolment Patterns (Repetition rates for three targeted provinces)	Used a balanced panel: With the availability of a learner unique identifier, Resep could link learner unit-level data from schools that submitted every year from 2016 to 2023 and followed the progression of the same cohort of learners.
LURITS	Enrolment Patterns (Repetition for South Africa) and Learner Mobility	Used a balanced panel for two consecutive years: Linked learner-unit data from schools that submitted in each pair of years, for example for 2018 and 2019 or for 2020 and 2021.
DDD	School-Based Assessments (SBAs)	Learner unit level subject data from 2016 to 2022. (Linking SBA performance over time makes it possible to start investigating SBAs across schools, and also how this influences learner outcomes and learner flows.)
DDD	Subjects Taught by Teachers	With the availability of a unique teacher identifier, Resep could link a teacher to the grade and subjects that they taught over time.
NSC data from DDD	SBA data linked to NSC outcomes	Individual NSC (matric) examination data per year matched to the SBA data through a unique identifier (anonymised SA ID)
NSC from DBE	NSC Performance over time	NSC data from 2008 to 2021
Masterlist of Schools for 2021	Integration of data sets and providing relevant details of schools	Uniquely identify each school in the country through a school identifier, generally called the "EMIS number"

1.2 Contents of Chapter 2: Learner flows between grades, school-based assessments and the NSC examination

In response to the outbreak of the Covid-19 pandemic, one of the measures implemented was to introduce more leniency in assessments to ensure fairness to students who could not attend classes regularly due to school closures or rotating timetables introduced to maintain social distancing.[2] As a result, curriculum coverage significantly declined in 2020 and to a lesser extent in 2021. This reduced the content that learners were tested on in assessments, leading to an improvement in school-based assessment results that made it easier for students to pass and thus progress to the next grade. There was a roughly 50% reduction in the repetition rates in grades 10 and 11 in 2020 when compared to 2019, before they again rose slightly in 2021 and again in 2022, but remained significantly lower than in 2019.

This decrease in repetition rates meant that more students were able to advance to higher grades and school dropout rates decreased significantly. Particularly, the historically high repetition rates in grade 10 that had discouraged many learners from persevering to grade 12 resulted in more students reaching matric and participating in the matriculation exam. This left Umalusi, the certification body, with difficult questions on how to adjust marks in individual subjects in a way that would both maintain standards whilst being fair to learners affected by pandemic-era learning losses.[3]

2 Hoadley, 2020.
3 Ardington et al., 2021; Van der Berg, Hoadley, et al., 2022.

Chapter 2 also contains an analysis on the relationship between marks achieved in some key subjects in school-based assessments (SBAs) and then later in the National Senior Certificate or matric examination. In addition, the chapter investigates the rise of poor schools (schools in Quintiles 1, 2 or 3) performing well above expectations.

1.3 Contents of Chapter 3: Learner mobility across schools and provinces

Chapter 3 delves into the complex dynamics of learner mobility across South African schools and provinces, particularly between 2020 and 2021. This analysis covers various dimensions of mobility, including mobility transitions within the same education levels, inter-level shifts from primary to secondary schools, geographic relocations across provinces, and transfers between public and private schools. The chapter employs LURITS data, utilising unique learner identifiers for accurate tracking of these movements.

Inter-school movements show the highest mobility rates in grade 10 and the lowest in grade 11. Repeaters are slightly more mobile than non-repeaters. Mobility between provinces appears quite consistent over time, with limited indication that the pandemic had a large influence, and with surprisingly low net migration to Gauteng. Cohort analysis of grade 8 learners in Gauteng from 2018 to 2022 also shows that few learners (about 2%) switch schools more than once in secondary school. A net shift from public to private schools during the pandemic occurred only in the two richest provinces, Gauteng and the Western Cape, supporting anecdotal evidence that financial pressure limited or even reversed flows to private schools. Visual representations, such as dot density and heat maps, further elucidate these mobility trends to offer an enhanced understanding of migration patterns within the education system.

1.4 Contents of Chapter 4: Teachers and the subjects they teach

Chapter 4 analyses educator data from SA-SAMS, focusing on public and independent schools in the Eastern Cape, Gauteng, and Limpopo from 2015 to 2023. The chapter explores the critical yet previously uncharted territory of what subjects educators specialize in at school. It addresses fundamental questions, such as the number of mathematics teachers in schools, which are essential for effective education planning. The chapter evaluates the completeness and consistency of the available educator data, noting a high degree of internal consistency but also identifying gaps in data for certain educator demographics, particularly for older educators and those not in permanent employment. This limitation is crucial as it affects the ability to calculate accurate attrition rates, a key factor for forecasting the demand for new teachers. The chapter underscores the significance of SA-SAMS as a comprehensive tool not just for building a learner database, but also for its potential in system-wide monitoring and planning of educator resources. This represents a pivotal step in harnessing SA-SAMS data for strategic education planning and provides a foundation for future inquiries into the effective management of human resources in schools.

CHAPTER 2
LEARNER FLOWS BETWEEN GRADES, SCHOOL-BASED ASSESSMENTS AND THE NSC EXAMINATION

Greater leniency is observed in some of the SBA results for the pandemic years that are discussed later in this report.

2.1 Greater leniency in response to the pandemic and its effects

During the pandemic and in reaction to the closure of schools, the education authorities considered that it would be fair to learners who had missed so much school through no fault of their own to be subject to more lenient promotion rules.[4]

For grades 10 and 11, the DBE increased the SBA component of the promotion requirements from 25% to 60%. Moreover, controlled tests replaced examinations at the end of the year, and it was required that these tests *"should only be set on content taught, content not taught cannot be assessed"*[5]

Adjustments were also made to the assessment rules for grades 4 to 11. The requirement for Home Language was reduced from 50% to 45%, that for the (usually English) First Additional Language from 40% to 35%. For Mathematics, the promotion requirement of 40% was dropped altogether, meaning that if all other requirements were met, Mathematics performance would be condoned, irrespective of performance. Moreover, it was explicitly stated that learners whose Mathematics grade 9 mark has been condoned would still be able to choose to continue with Mathematics in grade 10, i.e. they would not be compelled rather to take Mathematics Literacy.[6]

Greater leniency is observed in some of the SBA results for the pandemic years that are discussed later in this report.

4 Hoadley, 2020.
5 Circular S7, Revised promotion requirements for Grades 10 and 11 for the 2020 year.
6 National Assessment Circular 5 of 2022.

The leniency measures introduced to accommodate learners during the pandemic had an immediate effect on school-based assessments. The year 2020 marked a significant rise in the proportion of learners attaining pass levels across all subjects and provinces compared to 2019. Table 2.1 displays the percentage of learners in three provinces scoring 40% or above in 2019 and 2020 across four subjects. Particularly noteworthy is the surge in pass rates in Physical Science – in Limpopo, it increased from 17% in 2019 to 40% in 2020. English consistently demonstrates higher pass rates compared to other subjects, with even further rises of more than ten percentage points across all provinces.

TABLE 2.1 Percentage of Gr10 learners in three provinces achieving 40% or more in four subjects in Term 4 of 2019 and 2020

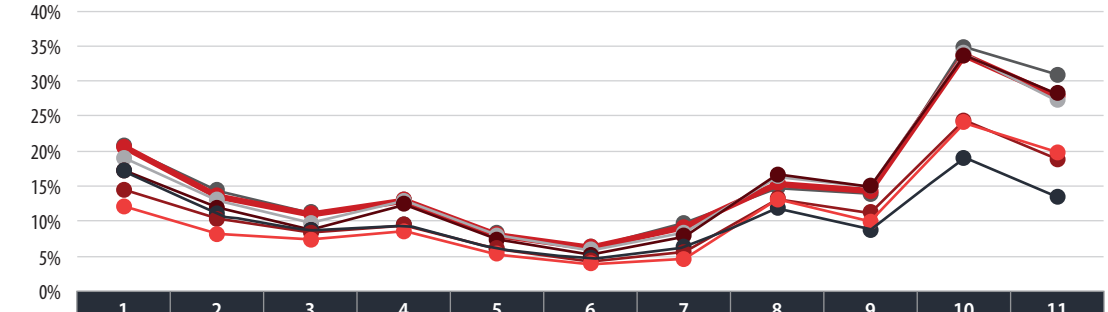
	Maths pass rate GR 10	Maths Lit pass rate GR10	Physical Science Gr 10	English First Additional Language (EFAL) pass rate Gr10
2019 EC	10%	16%	18%	70%
2020 EC	31%	36%	33%	83%
2019 GT	16%	24%	28%	69%
2020 GT	29%	36%	41%	80%
2019 LP	12%	15%	17%	72%
2020 LP	30%	33%	40%	85%

Source: Calculated from DDD data

2.2 Repetition trends and learner flows to higher grades

The greater leniency and resultant raised SBA marks led to a large decrease in repetition rates across the educational system. Figure 2.1, using data from the Eastern Cape from 2016 to 2021, demonstrates this change. Similar trends were observed in Gauteng and Limpopo. High school grades, particularly grades 10 and 11, saw substantial drops in repetition rates. In Gauteng, for example, the proportion of students repeating grade 10 fell from 33% in 2016 to 17% in 2020. Despite a slight increase in repetition rates in 2021 and 2022, they remained much lower than the levels seen before the pandemic.

FIGURE 2.1 Repetition rates by grade and year in the Eastern Cape, 2016–2022

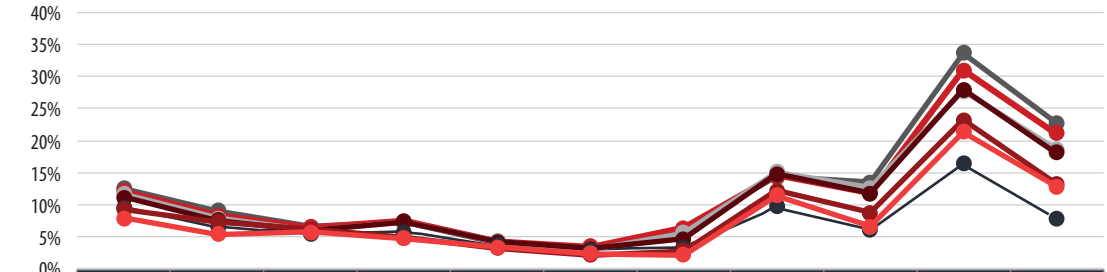


	1	2	3	4	5	6	7	8	9	10	11
2016	21%	14%	11%	13%	8%	6%	10%	15%	14%	35%	31%
2017	21%	14%	11%	13%	8%	6%	9%	15%	14%	34%	28%
2018	19%	13%	10%	13%	8%	6%	8%	16%	15%	34%	27%
2019	17%	12%	9%	12%	7%	5%	8%	17%	15%	34%	28%
2020	17%	11%	9%	9%	6%	5%	6%	12%	9%	19%	13%
2021	14%	10%	8%	9%	6%	4%	6%	13%	11%	24%	19%
2022	12%	8%	7%	8%	5%	4%	5%	13%	10%	24%	20%

Source: Calculated from DDD data

In Gauteng, the repetition rates also declined in 2020, especially in grade 10, where it decreased to 16% from 28% in 2019 (Figure 2.2). The rate increased to 23% in 2021 but slightly declined to 21% in 2022, still much lower than pre-pandemic levels.

FIGURE 2.2 Repetition rates by grade and year in Gauteng, 2016–2022

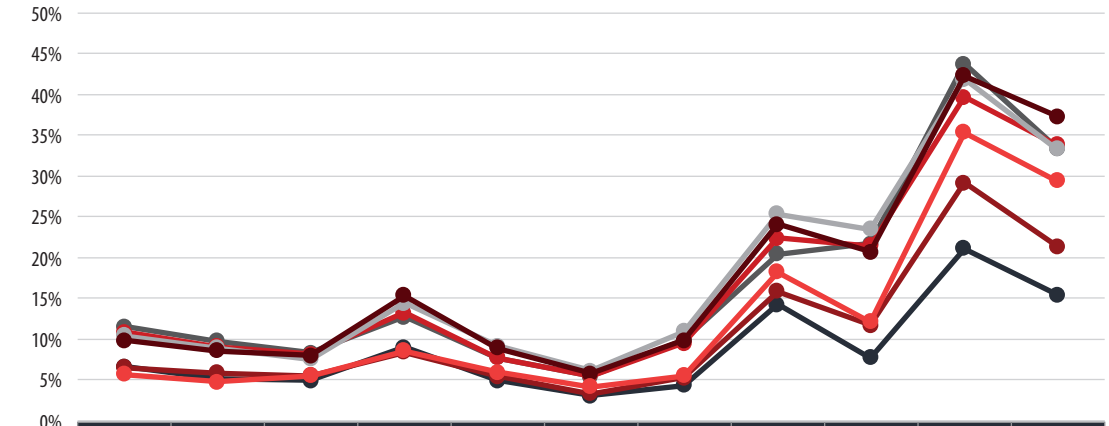


	1	2	3	4	5	6	7	8	9	10	11
2016	13%	9%	7%	7%	4%	3%	6%	14%	14%	34%	23%
2017	12%	8%	7%	7%	4%	3%	6%	15%	12%	31%	21%
2018	12%	8%	6%	7%	4%	3%	6%	15%	13%	28%	19%
2019	11%	8%	6%	7%	4%	3%	5%	15%	12%	28%	18%
2020	10%	6%	6%	6%	4%	3%	3%	10%	6%	16%	8%
2021	9%	7%	6%	5%	3%	2%	3%	12%	9%	23%	13%
2022	8%	5%	6%	5%	3%	2%	2%	11%	7%	21%	13%

Source: Calculated from DDD data

Limpopo mirrored this trend, with a marked drop in repetition rates in 2020 in grade 10 (Figure 2.3). In comparison to the other two provinces, Limpopo had the highest repetition rates in 2019, and these rates decreased from 42% to 21% in 2020. Since 2020, Limpopo has also experienced an increase in repetition rates at a much faster pace than the other provinces; it rose to 29% in 2021 and a staggeringly high 35% in 2022.

FIGURE 2.3 Repetition rates by grade and year in Limpopo, 2016–2022

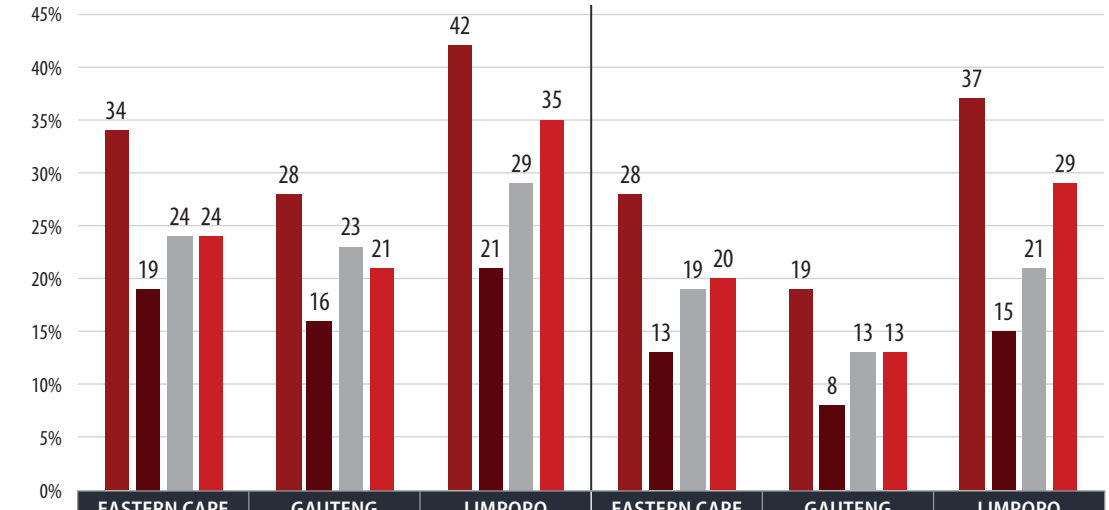


	1	2	3	4	5	6	7	8	9	10	11
2016	11%	10%	8%	13%	8%	5%	10%	20%	22%	44%	33%
2017	11%	9%	8%	13%	8%	5%	9%	22%	21%	40%	34%
2018	10%	9%	7%	14%	9%	6%	11%	25%	23%	42%	33%
2019	10%	9%	8%	15%	9%	6%	10%	24%	21%	42%	37%
2020	7%	5%	5%	9%	5%	3%	4%	14%	8%	21%	15%
2021	6%	6%	6%	8%	5%	3%	5%	16%	12%	29%	21%
2022	6%	5%	5%	9%	6%	4%	5%	18%	12%	35%	29%

Source: Calculated from DDD data

To compare the changing repetition rates across the three provinces, repetition rates are shown in Figure 2.4 for grades 10 and 11 for all three provinces from 2019 to 2022. Here the different levels of repetition across the provinces are clearly visible. What stands out particularly is the continuing rise in repetition rates in Limpopo in 2022, whilst in Eastern Cape and Gauteng repetition rates in both grade 10 and grade 11 have stabilised from 2021 to 2022.

FIGURE 2.4 Repetition rates by year and province for Gr10 & Gr11



	EASTERN CAPE	GAUTENG	LIMPOPO	EASTERN CAPE	GAUTENG	LIMPOPO
	Gr10			Gr11		
2019	34	28	42	28	19	37
2020	19	16	21	13	8	15
2021	24	23	29	19	13	21
2022	24	21	35	20	13	29

Source: Calculated from DDD data

Table 2.2 illustrates the changes in student progression patterns before and during the pandemic. Historically, there have been large rates of repetition or dropout in grade 10. In all three provinces shown, this still applied to the 2017 grade 10 cohort that reached matric before the pandemic. From 2020 there were declines in repetition and dropout, so it benefited the 2019 grade 10s only in grade 11 and the 2020 grade 10 cohorts in both 2020 and 2021. As a result, the proportion of students advancing to matric without repeating a grade or dropping out increased across all three provinces for the cohorts examined. In 2017 and 2019, 54% and 53% of grade students, respectively, progressed to grade 11 in the Eastern Cape. In 2020, this increased significantly, with 70% of grade 10 students moving to the next grade, due to the relaxed assessment criteria during the pandemic. This led to an increase in the proportion of grade 10 cohorts reaching matric without repetition or dropout, from 38% in 2017 to 52% in 2020 in the Eastern Cape, a rise of 14 percentage points. Gauteng and Limpopo experienced even larger increases of 17 and 21 percentage points respectively, with Limpopo starting from a lower base due to historically high repetition and dropout rates.

TABLE 2.2 Learner flow-through patterns without repetition from grade 10 to grade 12 for three grade 10 cohorts and across three provinces

	Eastern Cape			Gauteng			Limpopo		
	2017–19	2019–21	2020–22	2017–19	2019–21	2020–22	2017–19	2019–21	2020–22
Gr10	100%	100%	100%	100%	100%	100%	100%	100%	100%
Gr11	54%	53%	70%	56%	55%	71%	50%	48%	68%
Gr12	38%	44%	52%	40%	49%	57%	31%	39%	52%

Note: Calculated from a balanced panel of schools in the DDD data. The 'flow through' measured here is the proportion reaching the grade in question without any (further) repetition or dropping out.

Source: Calculated from DDD data

The 2021 matriculation results, while ostensibly adhering to traditional standards, saw a notable increase in candidates due to reduced repetition and dropout rates in 2020. The number of public school students writing matric grew by 41%, with uneven distribution across different quintiles, as Table 2.3 shows. Adjustments in mark distributions by Umalusi resulted in pass rates and bachelor's pass rates similar to pre-pandemic years, despite the increased cohort size and accumulated learning deficits.

TABLE 2.3 Increase in matric candidates from 2019 to 2021

	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5	Total
2019	102 720	104 609	113 550	65 566	95 737	**482 182**
2021	156 656	159 035	174 738	83 394	106 191	**680 014**
% increase	**53%**	**52%**	**54%**	**27%**	**11%**	**41%**

Note: Only schools for which quintile rankings are available in the matric results.

Source: NSC results

2.3 The choice between Mathematics and Mathematical Literacy in grade 10

In the Further Education and Training Phase, learners have to select whether to take Mathematics or Mathematic Literacy from grade 10. Mathematics is a more difficult subject than Maths Lit, but is an important gateway subject in that good performance in Mathematics is often a requirement for entering many university degree courses. This choice is thus an important one that affects later success at school and perhaps even whether learners pass matric, but also what future options they have in terms of university studies and careers.

The pandemic prompted relaxation in assessment rules, including curriculum adjustments, leading to increased learner promotions despite substantial learning losses. This raises concerns about learner readiness for advanced curricula and increases classroom heterogeneity.[7] The focus here as in Section 2.6 is on the predictive value of SBAs for grade progression, continuation to matric and performance in matric.

An essential decision for grade 10 learners involves choosing between Mathematics and Mathematical Literacy. This choice affects university study options for those electing to do Mathematics and performing well in it, but electing to take Mathematics rather than Mathematic Literacy for those performing less well in Mathematics in grade 9 may endanger their academic progress. The choice of Mathematics by low-achieving grade 9 students, significantly influences their academic performance in later years, including progressing to matric without repeating and matric performance. Students who opted for Mathematical Literacy generally achieved higher scores in this subject in the matric exams than their counterparts in Mathematics. This suggests a potential strategic advantage in selecting Mathematical Literacy for learners struggling with Mathematics.

Tables A1 to A3 in the Appendix contain regressions tables that offer further insights or at least pointers to aspects of educational dynamics. Table A1 shows Ordinary Least Squares (OLS) regressions that examine the factors associated with performance in Grade 10 Mathematics, for those who elected to do Maths rather than Mathematical Literacy. Key variables like the most recent Grade 9 Mathematics mark and whether the learner had repeated grade 9 show significant predictive power for grade 10 performance. These variables, along with provincial differences and quintile rankings, offer a quite detailed picture of the factors influencing learner success in this critical subject and the Further Education and Training Phase. Being overage does not seem to add much predictive power – in most years it is not statistically significant, while it is not clear what role gender plays, after considering all the other predictors.

The subsequent two sets of regressions employ logit models to investigate factors affecting students' likelihood of grade 10 repetition and the likelihood of remaining on track (i.e. without repetition from grade 10) and passing matric on their first attempt. They show the significant role of variables such as taking Mathematics in grade 10, the marks obtained in

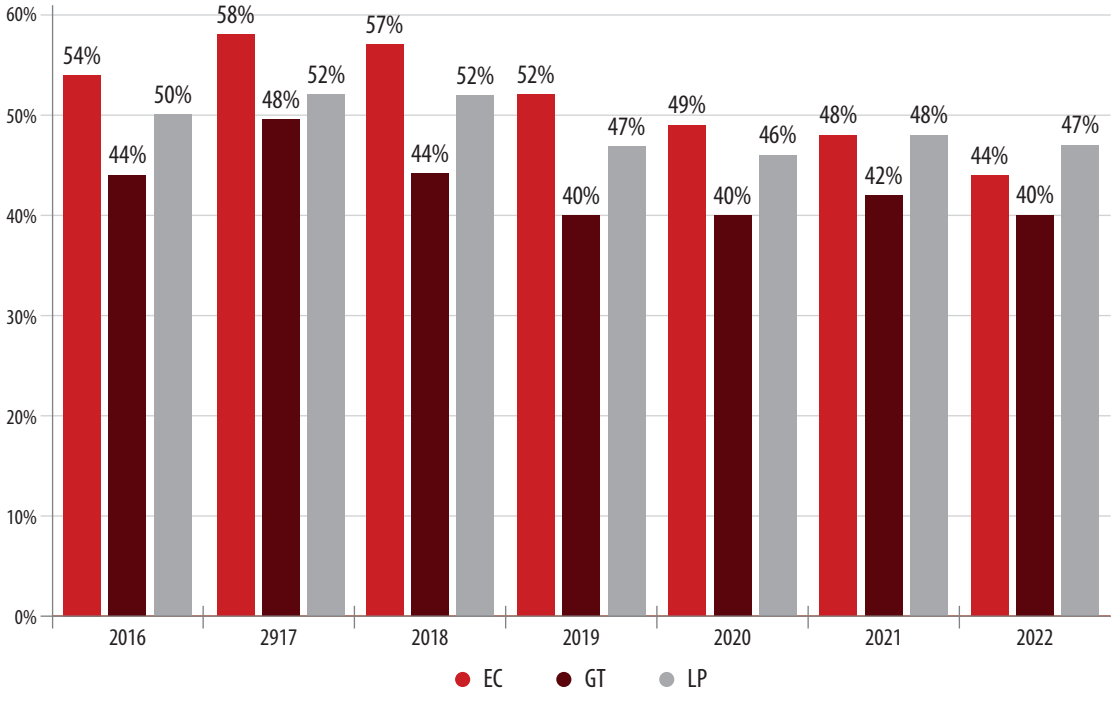
7 Wills & Hofmeyr, 2019.

Maths or Math Literacy, and the impact of repeating any year on the likelihood of academic success. Provincial and quintile differences again emerge as influential, along with gender and age, highlighting the complex interplay of various factors in learner outcomes. Compared to Gauteng, Limpopo learners are more likely to repeat Grade 10, while Eastern Cape learners are least likely to repeat the year, once all other factors in the model have been considered. In contrast, learners in both these provinces have a significantly lower likelihood of passing matric on track from grade 10 than their counterparts in Gauteng. Puzzlingly, the likelihood of passing matric without repetition from grade 10 appears to show opposite but significant signs for having elected to take mathematics in grade 10.

These regression offer one instrument for investigating the intricate dynamics between earlier performance, subject choice and academic success. However, much further work is necessary on this rich new data source.

Figure 2.5 shows that the percentage of new grade 10 learners taking Mathematics has been on a decreasing trend in the period 2016–2022. For example, while 54% of grade 10 learners in the Eastern Cape chose Mathematics in 2016, this declined to 44% in 2022. What is interesting is that a much smaller proportion of learners take Mathematics in Gauteng than in the other provinces. This is counter-intuitive: Gauteng generally performs better in international assessments than these two provinces and contains more learners from relatively wealthy households. Whereas almost three-quarters (73%) of Gauteng matriculants were from Quintile 4 and 5 schools, only 10% in the Eastern Cape and less than 6% in Limpopo are from those quintiles. But even amongst no-fee (Quintile 1 to 3) schools, around 45% of grade 10 learners elected to do Maths in 2022 as against only 34% in Gauteng.

FIGURE 2.5 Percentage of Gr10 learners electing to take Mathematics in three provinces, 2016–2022



Note: Not all schools submitted data in the earlier years, so earlier estimates for selecting Maths are less representative of the provinces concerned.

Source: Calculated from DDD data

For students that achieved a pass mark in Mathematics in grade 9 and then elected to do Mathematics rather than Maths Literacy in Gr10, the pass rate in 2020 group jumped by more than 20 percentage points in all three provinces. Interestingly, even students that failed grade 9 became more likely to pass grade 10 in 2020, although it still remained a minority of them who achieved a pass mark in grade 10.

GROWING PROMISE AMONG POOR SCHOOLS?

Historically, not many no-fee schools (Quintiles 1 to 3) have consistently performed well in terms of most learners reaching matric and a substantial share performing well in terms of gateway subjects, in particular Mathematics. But is this starting to change? A Resep report by Gustafsson (2016) was the first investigation of this matter; the same author has also produced another draft report that is nearing completion.

Using the DDD data, Quintile 1–3 schools were regarded as 'promising Maths schools' if more than a third of 2019 grade 9s were in grade 12 and elected to do Maths rather than Maths lit, and more than one-third of those learners performed achieved at least 60% in Maths in matric. Table 2.4 shows that there were 22 such schools in the Eastern Cape, 52 in Gauteng and 65 in Limpopo, confirming Gustafssson's observation that Limpopo was producing a relatively large number of promising schools. The table shows that a rising share of these schools matrics were doing Maths and obtaining 60% or more. This seems to support the evidence of a rising tide of promising schools.

Yet analysis of matric exam results for the same no-fee quintiles, or for black Africans, show no evidence of consistent growth in the four years before the pandemic. Does this simply mean that some promising schools are rising, while other poor schools that were perhaps performing well earlier are again falling behind? ***Identifying and supporting promising schools to sustain their rise may be an important policy imperative.***

TABLE 2.4 Percentage of learners achieving 60%+ marks in Mathematics in promising schools in three provinces, 2018–2022

	Eastern Cape	Gauteng	Limpopo
No. of schools	22	52	65
2018	9%	30%	10%
2019	10%	28%	9%
2020	14%	31%	12%
2021	20%	31%	16%
2022	29%	38%	28%

Source: Calculated from DDD data

To place this discussion within context, Table 2.5 below shows that 36% of matriculants wrote Mathematics in 2021. For the three provinces mainly considered in this publication, the proportion taking Maths was 47% in the Eastern Cape, 42% in Limpopo and only 31% in Gauteng.

The effect of the choice of the more difficult Mathematics rather than Mathematical Literacy is reflected in part in the percentage of those taking Maths achieving a C symbol (60%) or higher. In Gauteng this was 21%, as against only 11% in both Eastern Cape and Limpopo.

This analysis underscores the importance of careful consideration in subject choice at the grade 10 level, particularly between Mathematics and Mathematical Literacy, as it holds significant implications for academic trajectories and success in matriculation.

TABLE 2.5 Mathematics candidates and performance in the 2021 matric examination by province

	% taking Maths	% of all writing matric achieving 60% in Maths	Learners achieving 60% in Maths as % of those taking Maths
EC	47%	5%	11%
FS	36%	6%	17%
GP	31%	7%	21%
KZ	36%	6%	16%
LP	42%	5%	11%
MP	42%	6%	14%
NW	24%	4%	17%
NC	21%	3%	14%
WC	26%	7%	28%
SA	**36%**	**6%**	**16%**

Source: NSC examination data

2.4 Analysis of 2021 matric results

In 2021, over 10 000 learners wrote exams in 33 of the total 89 subjects. Table 2.6 presents these subjects along with their average marks. Mathematics and Technical Mathematics recorded the lowest average at 35%, with Geography and Economics also falling below 40%. The last column shows the matric failure rate for those candidates writing each subject, with those recording a matric pass mark of greater than 80% highlighted.

TABLE 2.6 Matric subjects that were written by more than 10 000 learners in 2021

	Candidates	Average	Failed matric
Life Orientation	703 822	61%	24%
English First Additional Language	583 025	54%	26%
Mathematical Literacy	433 793	42%	25%
Life Sciences	379 873	42%	25%
Geography	353 858	39%	26%
Mathematics	256 851	35%	22%
Business Studies	240 284	47%	22%
History	223 686	51%	25%
Tourism	197 168	55%	22%
Physical Sciences	195 561	40%	23%
IsiZulu Home Language	180 357	67%	25%
Economics	137 585	38%	25%
Agricultural Sciences	122 736	41%	30%
English Home Language	120 794	55%	13%
Accounting	105 049	44%	19%
IsiXhosa Home Language	103 269	68%	28%
Afrikaans First Additional Language	92 840	55%	12%
Sepedi Home Language	81 108	60%	33%
Setswana Home Language	56 472	61%	24%
Consumer Studies	48 270	48%	19%
Afrikaans Home Language	46 861	57%	15%
Computer Applications Technology	42 354	48%	12%
Sesotho Home Language	39 776	62%	20%
Engineering Graphic and Design	36 583	49%	16%
Xitsonga Home Language	32 003	62%	30%
SiSwati Home Language	21 442	66%	25%

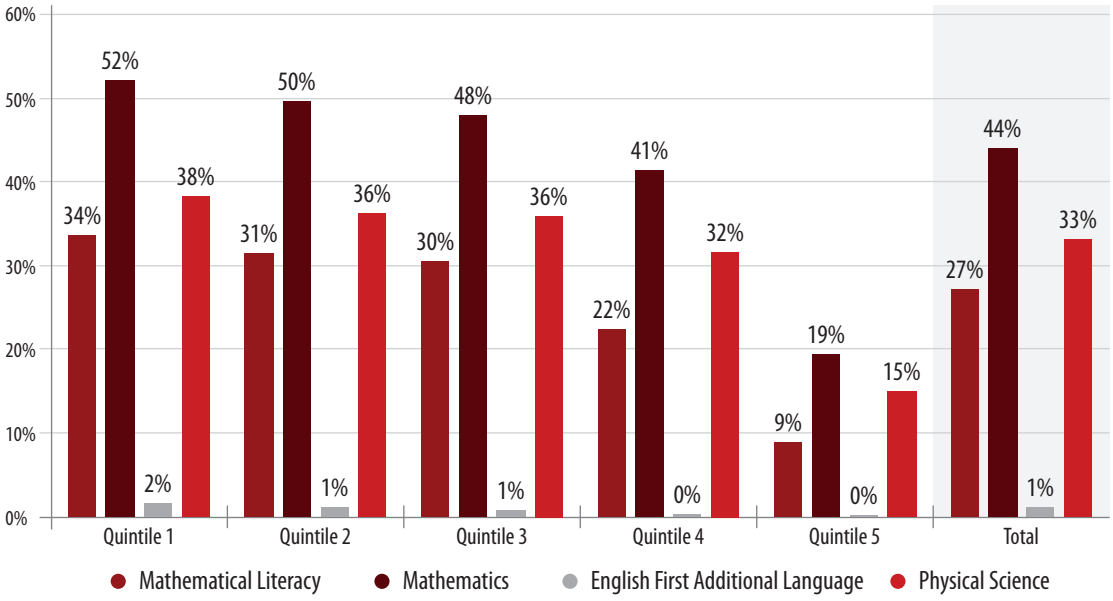
	Candidates	Average	Failed matric
Tshivenda Home Language	21 322	72%	29%
Dramatic Arts	17 994	57%	16%
IsiZulu First Additional Language	17 074	74%	17%
Afrikaans Second Additional Language	15 940	44%	21%
Religion Studies	14 489	45%	21%
Technical Sciences	14 368	42%	17%
Technical Mathematics	13 154	35%	17%
Total all subjects		**50%**	**24%**

Note: Highlighted cells in the middle column indicate the ten subjects on this list with the lowest average mark, while highlighted cells in the column on the right shows the selection of subjects chosen by those experiencing the lowest matric failure rate.

Source: NSC examination data

Figure 2.6 illustrates the percentage of candidates scoring below 30% in key subjects (the lowest pass mark).[8] At least a quarter of all candidates who wrote each of Mathematical Literacy, Mathematics and Physical Science failed to achieve a 30% mark in the respective subject in 2021. In contrast, English First Additional Language had a very low failure rate. There are large differences across the quintiles for the two Maths subjects in particular. Those learners achieving less than 30% for any of these subjects must then obtain at least 30% for all other subjects, in order to pass matric, and must also obtain 40% for their home language and two other subjects.

FIGURE 2.6 % of matric candidates scoring below 30% in four key subjects, South Africa 2021

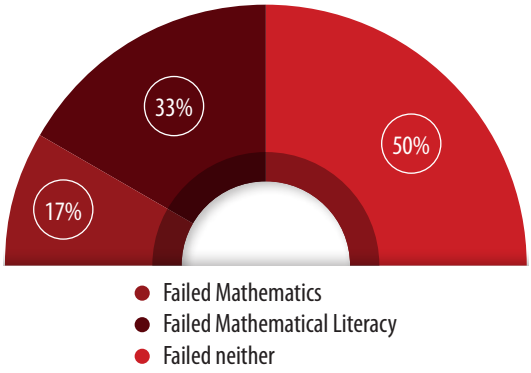


Source: NSC examination data

8 To get an NSC pass, learners must obtain at least 40% for their home language as well as for any two other subjects. In addition, for three further subjects out of the minimum of seven taken, learners must obtain at least 30%.

Poor performance in Mathematics or Mathematical Literacy is a feature of the matric results, and for many it also is closely related to whether they pass matric or not. Figure 2.7 delves into the relationship between failing matric and failing Mathematics or Mathematical Literacy. Approximately half of the 166 000 learners who failed matric had not passed Mathematical Literacy, a third of them had failed Mathematics, and 17% failed neither yet still did not pass the NSC. In contrast, of the 538 000 that passed matric, only 88 000 (16%) had failed either Maths or Maths Literacy.

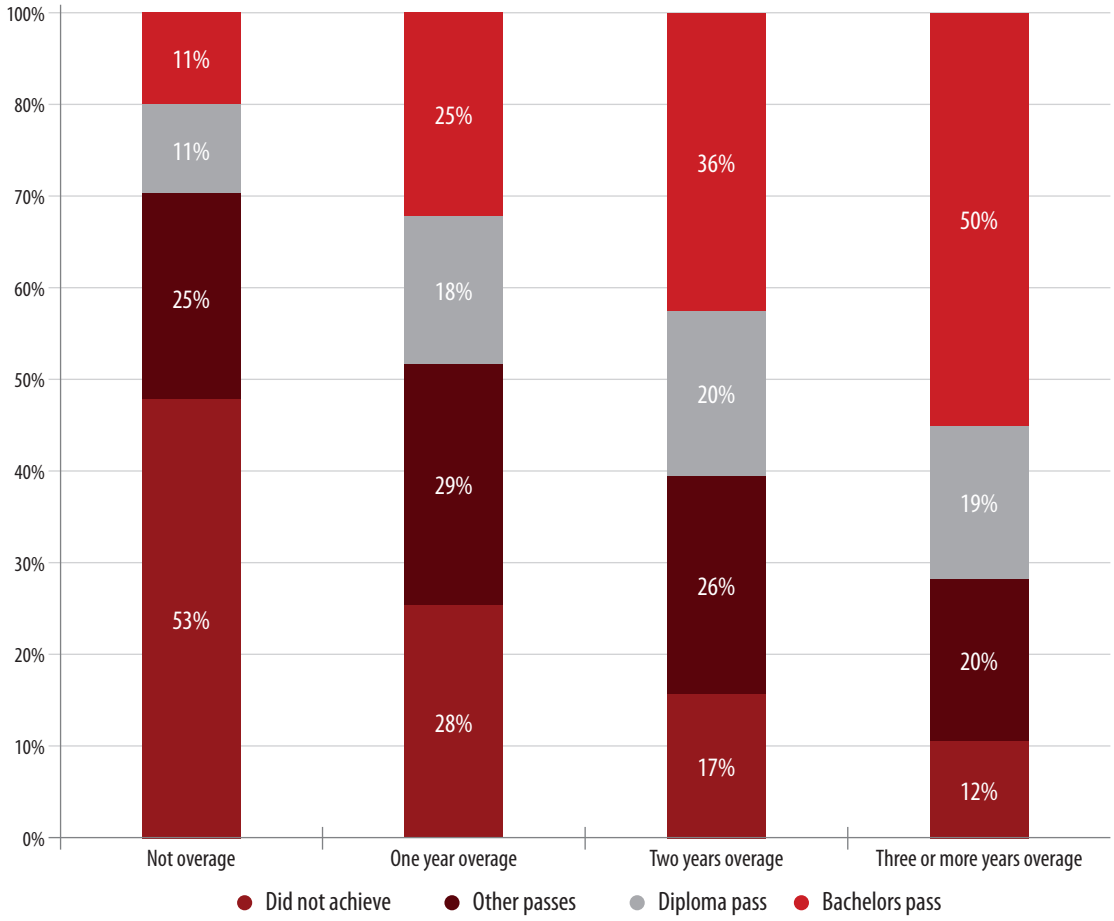
FIGURE 2.7 The association between failing matric and failing either Mathematics or Mathematical Literacy, South Africa 2021



Source: NSC examination data

Over age learners perform worse than learners who are in the appropriate age-for-grade, with learners who are three or more years over age performing particularly poorly, as Figure 2.8 shows.

FIGURE 2.8 Matric performance by years overage, South Africa 2021



Source: NSC examination data

2.5 How well do SBAs predict later school performance?

SBAs serve a dual purpose. As summative assessments, they determine learners' promotion to subsequent grades. Formatively, they provide crucial feedback, aiding learners in improving their performance. Given that SBAs are designed and implemented by individual schools' teachers, considerable variation in assessment standards is likely, especially in primary schools where feedback mechanisms are less developed. In contrast, high school assessments tend to align more closely with matric standards, offering a better predictive measure of matric performance. Of course, the fact that learners have different learning trajectories would also mean that assessments in later grades would better track the actual matric examination performance.

Table 2.7 illustrates the correlation between earlier grade SBA marks and the 2019 Eastern Cape matric cohort's final examination marks for four different subjects.[9] While correlations are moderate in lower grades, they significantly increase in grade 12, aligning closely with the NSC standards. This pattern may indicate variability in assessment quality at lower grades or in learning trajectories between grades.

TABLE 2.7 Correlation between NSC (matric) marks and SBA marks achieved in earlier grades by the 2019 Eastern Cape matric cohort in for four key subjects

	Maths	EFAL	MathsLit	Science
Gr9	0.43	0.52	n/a	n/a
Gr10	0.66	0.69	0.58	0.62
Gr11	0.81	0.75	0.74	0.73
Gr12	0.92	0.83	0.83	0.87

Source: Calculated from DDD data from Term 4 up to grade 12, and from Term 2 in grade 12. In lower grades. correlations with eventual NSC marks could only be calculated for those schools that submitted data for the relevant years. For example, fewer schools had submitted in the grade 9 cohort of 2016, so only those schools are considered for the correlation with the subsequent marks those learners achieved in the NSC examination in 2019.

Table 2.8 shows that the correlation between the matric examination marks in Mathematics and marks in earlier grades varies by quintiles, with generally higher correlations in Quintiles 4 and 5, but in grade 12 the differences are negligible.

Higher correlations in Quintiles 4 and 5 suggest better alignment with matric standards, though by grade 12, disparities across quintiles diminish. Similar trends are observed in other provinces.

9 Only learners who progressed from Gr9 to Gr12 without repetition were considered.

TABLE 2.8 Correlation between Mathematics NSC marks and SBA marks achieved in earlier grades by the 2019 Eastern Cape matric cohort by school quintile

	Q1	Q2	Q3	Q4	Q5
Gr9	0.39	0.39	0.42	0.34	0.55
Gr10	0.50	0.53	0.56	0.71	0.86
Gr11	0.73	0.68	0.77	0.92	0.93
Gr12	0.90	0.89	0.91	0.94	0.92

Source: Calculated from DDD data from Term 4 up to grade 12, and from Term 2 in grade 12. In lower grades. correlations with eventual NSC marks could only be calculated for those schools that submitted data for the relevant years. For example, fewer schools had submitted in the grade 9 cohort of 2016, so only those schools are considered for the correlation with the subsequent marks those learners achieved in the NSC examination in 2019.

When one considers gender, it appears that the correlation between earlier and matric exam marks is higher for girls than boys. The is illustrated in Table 2.8 for the Eastern Cape, with similar trends in Limpopo and Gauteng. This may perhaps be accounted for by greater variation in the learning trajectories of boys than girls.

Analysing gender differences, Table 2.9 indicates higher correlations between SBA and NSC marks for girls compared to boys in the Eastern Cape, a trend consistent across all three provinces. This disparity might suggest greater variability in boys' learning trajectories.

TABLE 2.9 Correlations for four subjects between grade 10 SBA results and subsequent NSC results for the 2019 Eastern Cape matric cohort

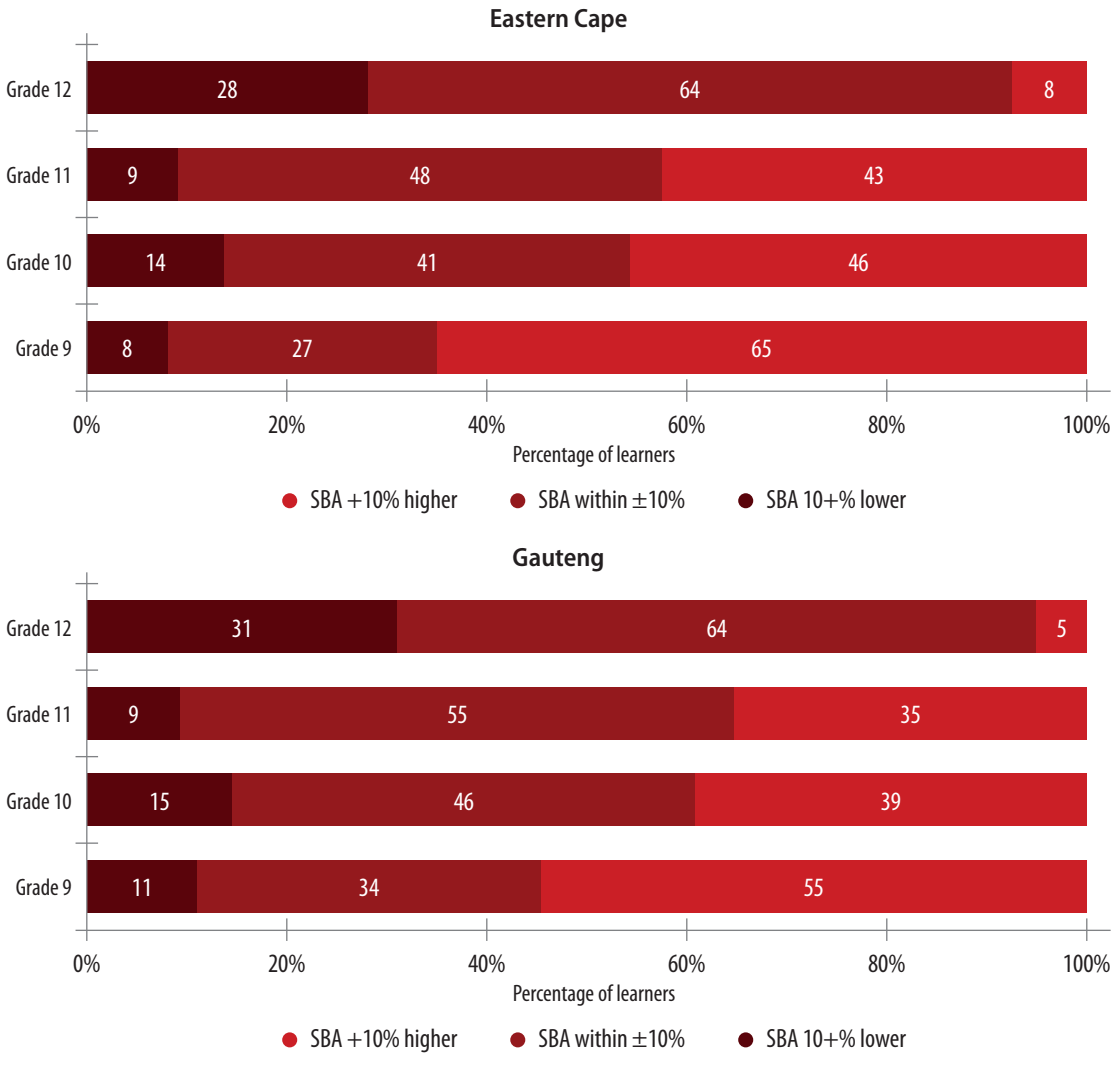
	Girls	Boys
Maths	0.70	0.63
EFAL	0.70	0.66
MathsLit	0.63	0.54
Science	0.66	0.59

Source: Based on the DDD data. For lower grades, correlations could only be calculated for those schools that submitted data in that year.

Figures 2.9 A, B and C illustrate the relationship between Mathematics NSC marks and marks in earlier grades by the same learners. Typically, most learners score significantly higher marks in the SBAs than in the matric exams, especially in the lower grades. A notable feature is the much lower marks in grade 9 compared to the matric exam, indicating that SBAs in lower grades may overestimate learner capabilities compared to the appropriate standard. For example, in the Eastern Cape, 65% of grade 9 learners score at least 10 percentage points higher in SBAs than what they achieve in grade 12 NSC exams, while only 8% achieve at least 10 percentage points lower marks than they eventually achieve in the matric exam. The pattern is consistent across all three provinces examined. As students progress to grades 10 and 11, the gap between their SBA marks and what they later achieve in the NSC narrows.

An interesting shift occurs in grade 12, where SBA marks (in Term 2) are generally *lower* than those in the matric examination. In the Eastern Cape, for instance, a higher proportion of learners (28%) are likely to see an increase of at least 10 marks in the matric exam compared to those (8%) who might experience a similar decrease. Gauteng exhibits a similar pattern, whereas in Limpopo those who gain or lose at least 10 marks are roughly in balance.

FIGURE 2.9 A, B AND C Distribution of the difference between the SBA marks in different grades and the eventual marks obtained in the NSC in Mathematics for the 2022 matric cohort



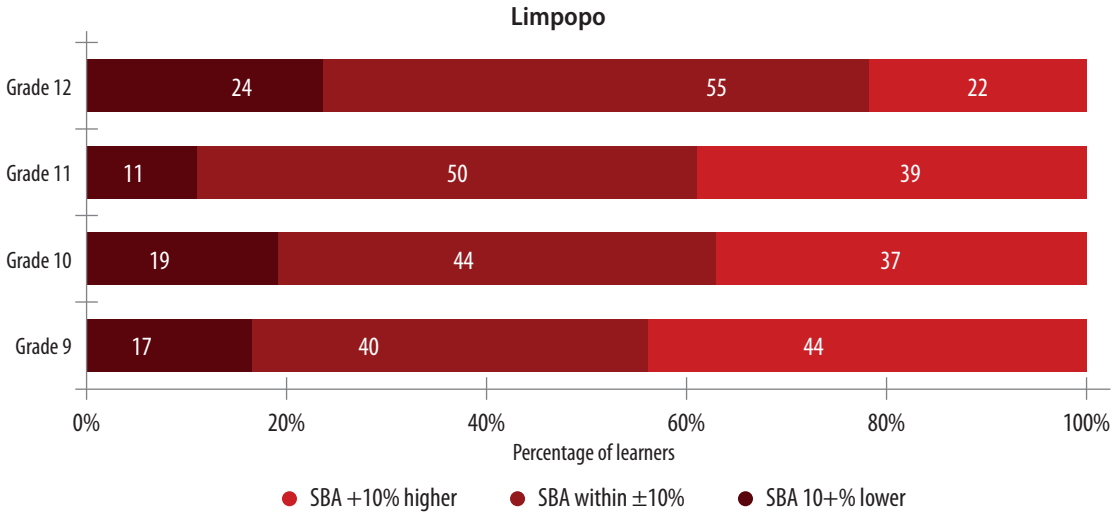
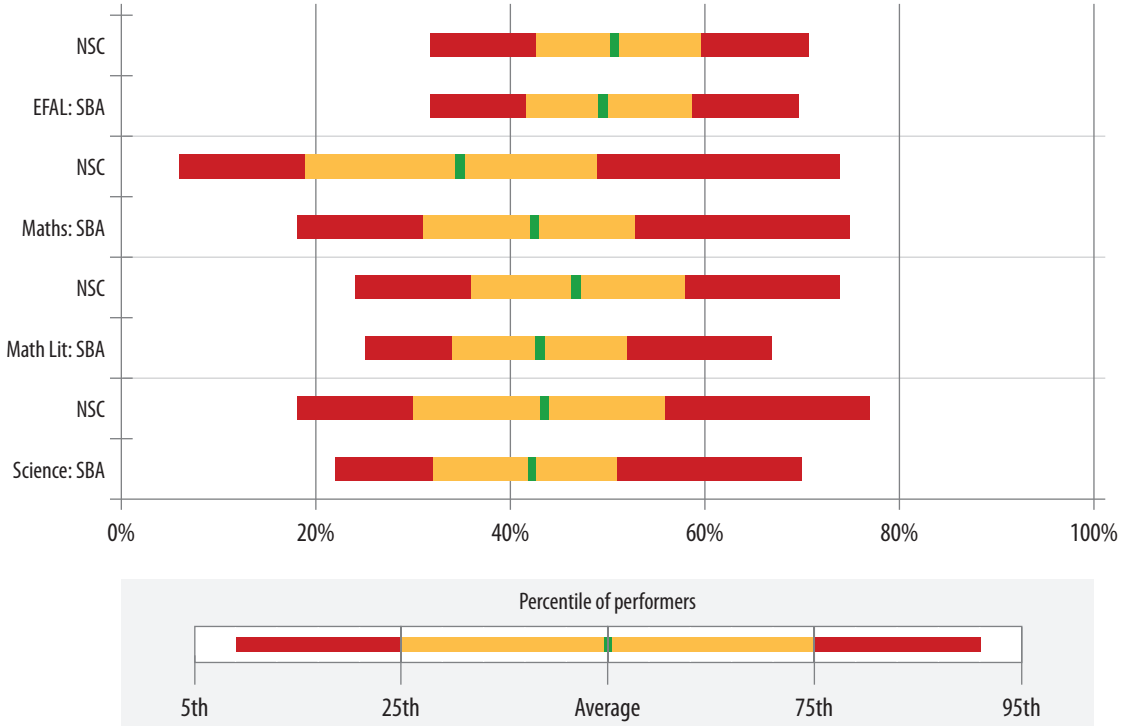


Figure 2.10 compares the distribution of SBA and NSC marks across four subjects in the Eastern Cape. A striking similarity that comes to the fore here is that the distribution of marks in English First Additional Language in grade 10 is quite similar to the distribution of the marks the same learners achieved when they wrote matric. For other subjects, however, the NSC marks show a far greater spread than is the case for the NSC marks.

FIGURE 2.10 Distribution of grade 10 SBA marks and eventual NSC marks achieved in 2022 by the same learners in the Eastern Cape



In summary, these figures highlight an overarching trend of higher SBA marks in the earlier grades, with a convergence of SBA and NSC scores as students approach their final examinations. The data underscores the variability of academic performance across different stages of education and across provinces.

A further analysis carried out on the data was to determine the share of the variance in scores that can be described as differences between rather than within schools, or the so-called intra-class correlation coefficient or ICC. This is shown in Table 2.10 for Maths and EFAL, mainly for the 2019 matric cohort (the 2022 Eastern Cape cohorts is also shown for Mathematics). A large intra-class correlation coefficient is an indication of large performance differences between schools, while the within-school performance is highly correlated. South African intra-class correlation coefficients in international assessments have been shown to be very high in international comparison.

One would expect that differences in standards that different teachers or schools apply in the SBAs may lead to somewhat lower ICCs, as some of the differences between schools would then be muted. In contrast, the NSC differences, where a common standard applies, is perhaps a better measure of the inequality in performance within the school system. On the other hand, because of differential dropout, the matric mark may also not fully reflect the degree of inequality between schools.

There is no universal subject taken by all matriculants, however, so that there is some selection at play with whatever subject is considered. Mathematics is more often selected as a subject in matric by stronger students; in contrast, EFAL would not be taken by strong English home language students, thus perhaps under-emphasising inter-school differences.

Also surprising is that the ICC does not appear to be lower when more affluent Quintile 4 and 5 schools are excluded. On the other hand, when considering the final part of this table, it is clear that inequality grows across the grades in Mathematics in all three provinces considered, but that there are no similar patterns for EFAL.

What the ICCs shown here do indicate is that inequality between schools in South Africa is extraordinarily high. Far more analysis is necessary to fully understand how ICCs are influenced by varying SBA standards across schools. The fact that such data is now available for the first time makes this an important area for further research.

TABLE 2.10 Intra-class correlation coefficient in Mathematics and English First Additional Language for the 2019 matric cohort

	EC	*EC*	GT	LP	EC	GT	LP
	Maths	*Maths*	Maths	Maths	EFAL	EFAL	EFAL
	2019	***2022***	2019	2019	2019	2019	2019
National Senior Certificate (matric)							
All Schools	0.70	*0.75*	0.74	0.75	0.69	0.77	0.72
Q1–3	0.77	*0.80*	0.76	0.77	0.74	0.82	0.77
Q4–5	0.76	*0.80*	0.75	0.88	0.77	0.73	0.74
Gr10 SBAs							
All Schools	0.56	*0.63*	0.61	0.69	0.73	0.71	0.68
Q1–3	0.61	*0.67*	0.65	0.72	0.75	0.79	0.69
Q4–5	0.67	*0.81*	0.60	0.69	0.76	0.64	0.77
SBAs for different grades experienced by the 2019 matric cohort							
Gr9	0.57	*0.60*	0.56	0.58	0.65	0.76	0.63
Gr10	0.56	*0.63*	0.61	0.69	0.73	0.71	0.68
Gr11	0.65	*0.69*	0.70	0.76	0.69	0.74	0.67
Gr12	0.69	*0.73*	0.77	0.76	0.69	0.75	0.63

Note: The inter-class correlation coefficient expresses the variance between schools as a share of the total variance in marks on any given test or examination. It is thus a measure of the extent of inequality in performance between rather than within schools.

Source: Calculated from DDD data

CHAPTER 3
LEARNER MOBILITY

> This section delves into the dynamics of learner mobility from 2020 to 2021, capturing a wide spectrum of transitions.

3.1 Measuring movement of learners between schools

This section delves into the dynamics of learner mobility from 2020 to 2021, capturing a wide spectrum of transitions. These transitions encompass movements within the same educational level (for example, relocating from one primary school to another within the same region), as well as shifts between different educational levels (such as advancing from primary to secondary school between grades 7 and 8). Additionally, we consider geographic relocations across provinces and transitions between public and private (independent) schools. Factors influencing mobility range from geographic relocation and parental choice influenced by school performance to economic constraints and language considerations.

To accurately track these movements, we utilized the LURITS data of 2020 and linked it to the 2021 records using a unique learner identifier[10] assigned to each learner. Accurate and consistent use of this identifier is crucial for reliable tracking. The primary metric was the proportion of learners switching schools in each grade, relative to the total number of learners per grade within a province, as detailed in Table 3.1.

The data reveals significant insights into school mobility patterns. To show the extent to which learners switch schools between 2020 and 2021 in each grade,[11] Figure 3.1 shows the percentage of learners that switch schools in each grade between 2020 and 2021 **within** each province. The highest mobility rates occur in grade 10, with national averages around 12% (see Figure 3.1), peaking at 22% in the Eastern Cape and 15% in Mpumalanga and the Northern Cape. Grade 11 experiences the least mobility.

10 A unique identifier is a single, non-duplicated number that is assigned to, and remains with, a learner throughout his or her education career irrespective of whether the learner changes schools.

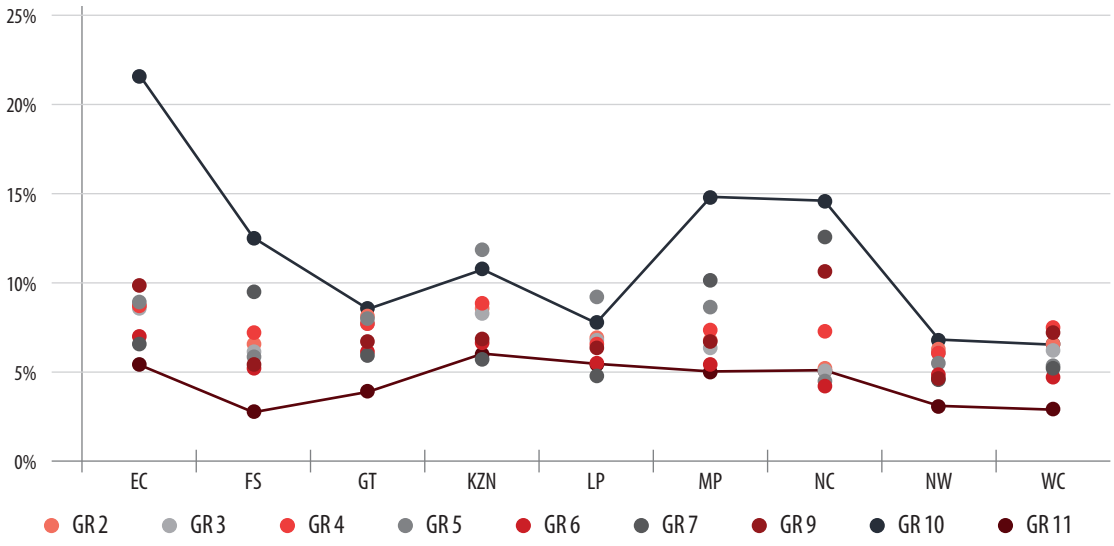
11 Grade 8 is omitted, as the transition from primary to secondary school typically occurs between grades 7 and 8.

TABLE 3.1 Learners switching schools from 2020 to 2021 as percentage of total learners in a grade in 2020 by province and grade

	Gr2	Gr3	Gr4	Gr5	Gr6	Gr7	Gr9	Gr10	Gr11
EC	8.7%	8.5%	8.7%	8.9%	7.0%	6.5%	9.9%	21.6%	5.4%
FS	6.5%	6.2%	7.2%	5.8%	5.2%	9.5%	5.4%	12.5%	2.7%
GT	8.1%	7.7%	7.7%	8.0%	6.1%	5.9%	6.7%	8.5%	3.9%
KZN	8.5%	8.3%	8.8%	11.9%	6.6%	5.7%	6.9%	10.7%	6.0%
LP	6.9%	6.8%	6.6%	9.2%	5.5%	4.7%	6.4%	7.8%	5.4%
MP	6.8%	6.3%	7.3%	8.6%	5.4%	10.2%	6.7%	14.8%	5.0%
NC	5.2%	5.1%	7.3%	4.5%	4.2%	12.6%	10.6%	14.6%	5.1%
NW	6.2%	6.1%	6.0%	5.5%	4.9%	4.6%	4.6%	6.8%	3.1%
WC	6.5%	6.2%	7.5%	5.3%	4.7%	5.2%	7.2%	6.5%	2.9%
SA	**8.0%**	**7.0%**	**8.0%**	**9.0%**	**6.0%**	**6.0%**	**7.0%**	**12.0%**	**5.0%**

Source: Calculated from LURITS data

FIGURE 3.1 Learners switching schools as a percentage of total learners in each grade by province and grade, 2020 to 2021



Source: Calculated from LURITS data

3.2 Mobility of repeaters

Repetition, or retention as it is often termed in developed countries, involves holding back learners who have not mastered the curriculum and, consequently, do not meet certain academic standards while their peers are promoted to the next grade.[12] There is a growing literature on repetition, both within South Africa and internationally.[13] Here there is a specific focus on learners who repeated a grade between 2020 and 2021, and whether they also switched schools, possibly to another province. The data, as outlined in Table 3.2, shows an intriguing pattern: a significant portion of mobility among repeaters occurred during the primary school phase, with the least mobility observed in grades 10 and 11. Notably, 15% of repeaters in grades 1, 3 and 6 switched schools.

TABLE 3.2 Mobility amongst repeaters by province and grade, 2020–2021

	EC	FS	GT	KZN	LP	MP	NC	NW	WC	Average
Gr11	12.5%	5.7%	9.8%	11.4%	9.2%	8.2%	4.9%	5.6%	5.9%	8.1%
Gr10	11.9%	4.5%	6.8%	12.6%	8.4%	7.8%	4.6%	4.8%	6.1%	7.5%
Gr9	14.8%	9.4%	11.8%	15.0%	10.9%	9.8%	7.0%	7.2%	9.3%	10.6%
Gr8	12.1%	6.5%	10.6%	13.9%	9.6%	8.9%	7.7%	5.5%	16.4%	10.2%
Gr7	16.7%	10.1%	19.1%	18.6%	11.3%	18.1%	7.0%	12.5%	17.1%	14.5%
Gr6	16.2%	14.6%	20.8%	20.8%	11.3%	13.2%	7.9%	14.1%	15.0%	14.9%
Gr5	15.9%	12.2%	19.8%	20.1%	11.4%	12.8%	8.0%	11.7%	13.9%	14.0%
Gr4	14.2%	10.1%	16.4%	16.8%	9.6%	12.2%	6.7%	9.5%	10.9%	11.8%
Gr3	17.0%	13.7%	18.3%	20.5%	13.6%	13.1%	8.2%	12.5%	15.0%	14.7%
Gr2	15.7%	13.9%	17.8%	19.5%	13.9%	13.4%	8.0%	12.8%	14.0%	14.4%
Gr1	15.9%	16.3%	18.2%	18.7%	14.2%	15.2%	8.7%	13.7%	16.1%	15.2%

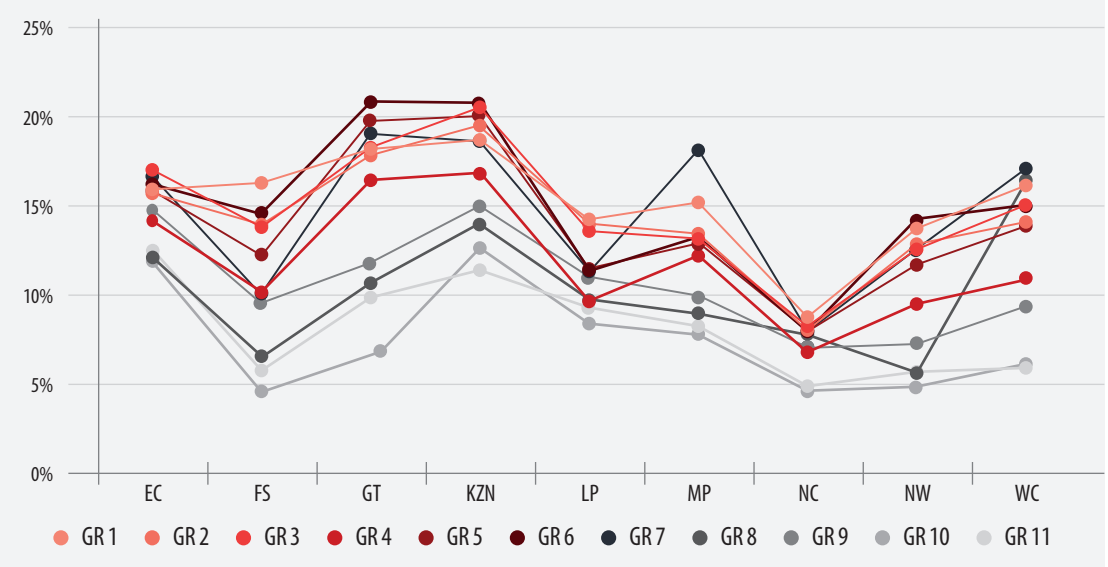
Source: Calculated from LURITS data

12 Ndaruhutse, Brannelly, Latham & Penson, 2008.

13 Branson & Lam, 2009; Brophy, 2006; DBE, 2012; Kabay, 2016; Kika & Kotze, 2019; Ndaruhutse, Brannelly, Latham & Penson, 2008; Van der Berg, Wills, Selkirk, Adams, Van Wyk, 2019; Van der Berg, Van Wyk, Selkirk, Rich & Deghaye, 2019; Weatherholt, Crouch, Jordan, Healey, Merseth & Eileen Dombrowski, 2018.

Figure 3.2 shows the percentage of learners who repeated in grades 2 to 11 between 2020 and 2021 who had also switched schools in each province, i.e. they repeated the grade in another school. Most mobility amongst repeaters takes place in the primary school phase, with the least in grades 10 and 11.

FIGURE 3.2 School mobility amongst repeaters by province and grade, 2020–2021



Source: Calculated from LURITS data

3.3 Inter-provincial learner mobility

Anecdotal evidence and findings from the NIDS-CRAM study suggested that distinct mobility patterns emerged during the pandemic. Using LURITS data spanning from 2018 to 2021 makes it possible to observe inter-provincial learner mobility between provinces before and during the Covid-19 pandemic. For accurate tracking, each learner was assigned a unique identifier, enabling data linkage across years. The provision of anonymised data provided by the DBE ensured privacy compliance.

Tables 3.3, 3.4 and 3.5 illustrate the out-of-province and into-province movements between 2020 and 2021. For instance, in 2020, 13 136 learners from the Eastern Cape moved to the Western Cape, and 13 812 learners relocated from Limpopo to Gauteng as indicated by the radar chart. Conversely, 15 383 learners moved from Gauteng to Limpopo.

Regarding **net mobility**, the overall result shows that 55 282 (3%) learners moved out of Gauteng while 49 889 learners moved into Gauteng in 2021, indicating quite limited net mobility of 6 393 to Gauteng between these two Covid years. Similar explanation applies for the rest of the provinces. The low net migration to Gauteng is surprising. A DBE report finds slightly different numbers,[14] but the same underlying pattern:

What is remarkable is that with the onset of the pandemic, Gauteng switched from being a major receiver of migrating learners to a major sender of such learners. This would be in line with anecdotal evidence that in response to the lockdowns of the pandemic, but also economic hardship, households in Gauteng sent learners to neighbouring provinces to stay with family. The evidence here suggests that a fair proportion of these learners ended up enrolled in schools in the receiving provinces.

Comparing across these three tables, movement between provinces was relatively consistent over time. For example, there was high mobility to and out of Gauteng between the years considered.

Some visual representations further elucidate mobility trends Figure 3.3 indicates significant inward-mobility from Limpopo, Kwazulu-Natal and the Eastern Cape into Gauteng. Conversely, Figure 3.4 shows that there was high outward-mobility from Gauteng to other provinces such as Limpopo and Kwazulu-Natal. Additionally, Figure 3.5 focuses on the substantial migration from the Eastern Cape to the Western Cape.

14 The DBE report was based on the same underlying Lurits data, but made more specific assumptions as to how to deal with unidentified in- and outflows than was done here.

TABLE 3.3 Mobility of learners across provincial boundaries, 2020–2021

	PROVINCE 2021											
Province 2020	EC	FS	GT	KZN	LP	MP	NC	NW	WC	TOTAL	Out of province mobility	% Of total out of province mobility
EC	1 608 235	1 101	7 939	6 721	267	677	271	1 194	13 136	**1 639 541**	**31 306**	2%
FS	1 015	651 815	3 955	692	293	496	500	1 058	375	**660 199**	**8 384**	1%
GT	6 916	4 477	2 141 295	10 959	15 383	7 054	742	7 392	2 359	**2 196 577**	**55 282**	3%
KZN	5 536	848	10 400	2 562 263	594	2 732	110	463	639	**2 583 585**	**21 322**	1%
LP	321	293	13 812	546	1 595 430	3 724	96	1 599	215	**1 616 036**	**20 606**	1%
MP	557	487	6 028	2 328	3 233	997 467	113	551	216	**1 010 980**	**13 513**	1%
NC	249	532	649	159	113	99	272 911	1 311	656	**276 679**	**3 768**	1%
NW	969	1 062	6 134	479	1 380	487	1 309	771 040	237	**783 097**	**12 057**	2%
WC	5 034	287	972	377	100	104	550	161	1 118 483	**1 126 068**	**7 585**	1%
Total	**1 628 832**	**660 902**	**2 191 184**	**2 584 524**	**1 616 793**	**1 012 840**	**276 602**	**784 769**	**1 136 316**			
Into province mobility	**20 597**	**9 087**	**49 889**	**22 638**	**21 363**	**15 373**	**3 691**	**13 729**	**17 833**			
% of total into province mobility	1%	1%	2%	1%	1%	2%	1%	2%	2%			

TABLE 3.4 Mobility of learners across provincial boundaries, 2019–2020

	Province in 2020												
	Province	EC	FS	GT	KZN	LP	MP	NC	NW	WC	Total	Out	% Out as Total
Province in 2019	EC	1 625 078	1 056	7 110	6 732	452	733	354	1 245	13 515	**1 656 275**	**31 197**	2%
	FS	823	643 051	4 067	752	342	566	571	1 212	389	**651 773**	**8 722**	1%
	GT	4 515	3 217	2 106 433	7 409	11 353	6 285	743	5 550	1 853	**2 147 358**	**40 925**	2%
	KZN	4 358	729	9 917	2 551 774	584	3 067	138	349	614	**2 571 530**	**19 756**	1%
	LP	213	268	15 928	506	1 551 124	5 335	138	1 564	191	**1 575 267**	**24 143**	2%
	MP	528	434	6 321	2 004	3 099	969 489	118	487	226	**982 706**	**13 217**	1%
	NC	276	525	663	100	120	114	269 353	1 503	835	**273 489**	**4 136**	2%
	NW	775	1 001	7 629	424	1 616	680	1 872	764 187	262	**778 446**	**14 259**	2%
	WC	2 618	195	1 400	309	117	244	893	146	1 084 419	**1 090 341**	**5 922**	1%
	Total	**1 639 184**	**650 476**	**2 159 468**	**2 570 010**	**1 568 807**	**986 513**	**274 180**	**776 243**	**1 102 304**			
	In	**14 106**	**7 425**	**53 035**	**18 236**	**17 683**	**17 024**	**4 827**	**12 056**	**17 885**			
	% In as Total	1%	1%	2%	1%	1%	2%	2%	2%	2%			

TABLE 3.5 Mobility of learners across provincial boundaries, 2018–2019

	PROVINCE IN 20219												
	Province	EC	FS	GT	KZN	LP	MP	NC	NW	WC	Total	Out	% Out
Province in 2018	EC	1 616 381	1 391	8 464	7 737	342	780	331	1 434	15 503	**1 652 363**	**35 982**	2%
	FS	844	596 794	4 044	705	257	573	577	1 089	327	**605 210**	**8 416**	1%
	GT	5 839	3 604	2 004 244	8 806	10 383	6 317	700	6 267	2 129	**2 048 289**	**43 655**	2%
	KZN	5 112	1 007	10 375	2 420 506	432	2 945	94	389	711	**2 441 571**	**21 065**	1%
	LP	320	336	13 976	525	1 544 118	4 011	88	1 547	174	**1 565 095**	**20 977**	1%
	MP	634	514	6 498	2 252	2 879	896 573	102	566	227	**910 245**	**13 672**	2%
	NC	333	633	879	113	91	110	262 052	1 624	917	**266 752**	**4 700**	2%
	NW	1 012	1 148	7 943	545	1 469	685	1 858	741 748	302	**756 710**	**14 962**	2%
	WC	7 421	215	1 276	440	64	107	750	136	989 101	**999 510**	**10 409**	1%
	Total	**1 637 896**	**605 642**	**2 057 699**	**2 441 629**	**1 560 035**	**912 101**	**266 552**	**754 800**	**1 009 391**			
	In	**21 515**	**8 848**	**53 455**	**21 123**	**15 917**	**15 528**	**4 500**	**13 052**	**20 290**			
	% In	1%	1%	3%	1%	1%	2%	2%	2%	2%			

FIGURE 3.3 Inward-mobility of learners from other provinces into Gauteng

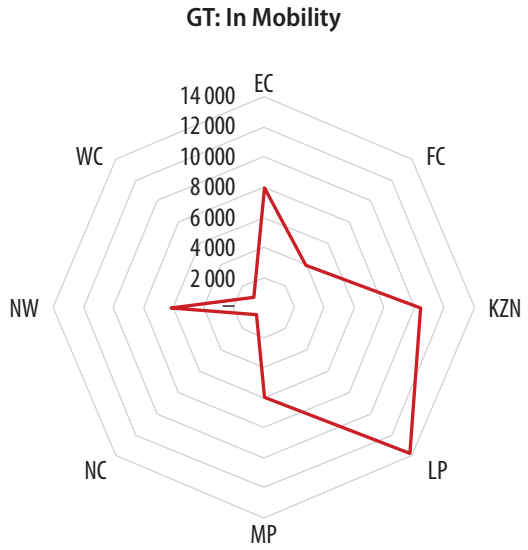


FIGURE 3.4 Outward-mobility of learners from Gauteng to other provinces

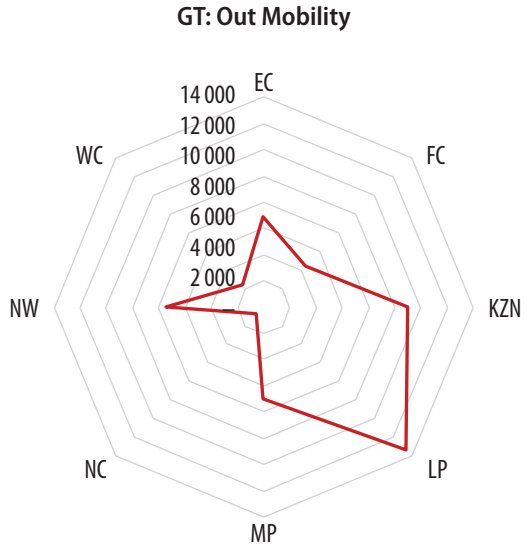
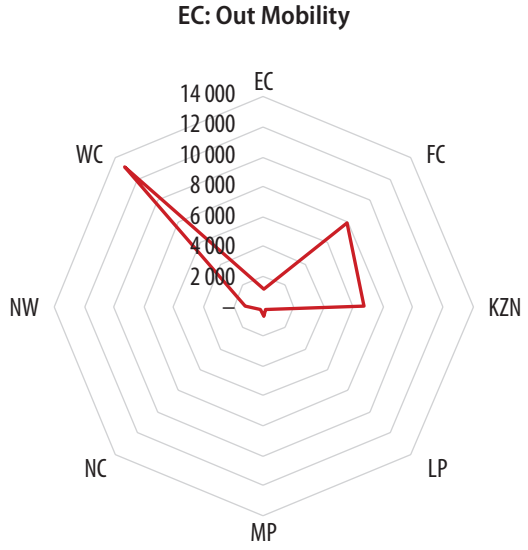


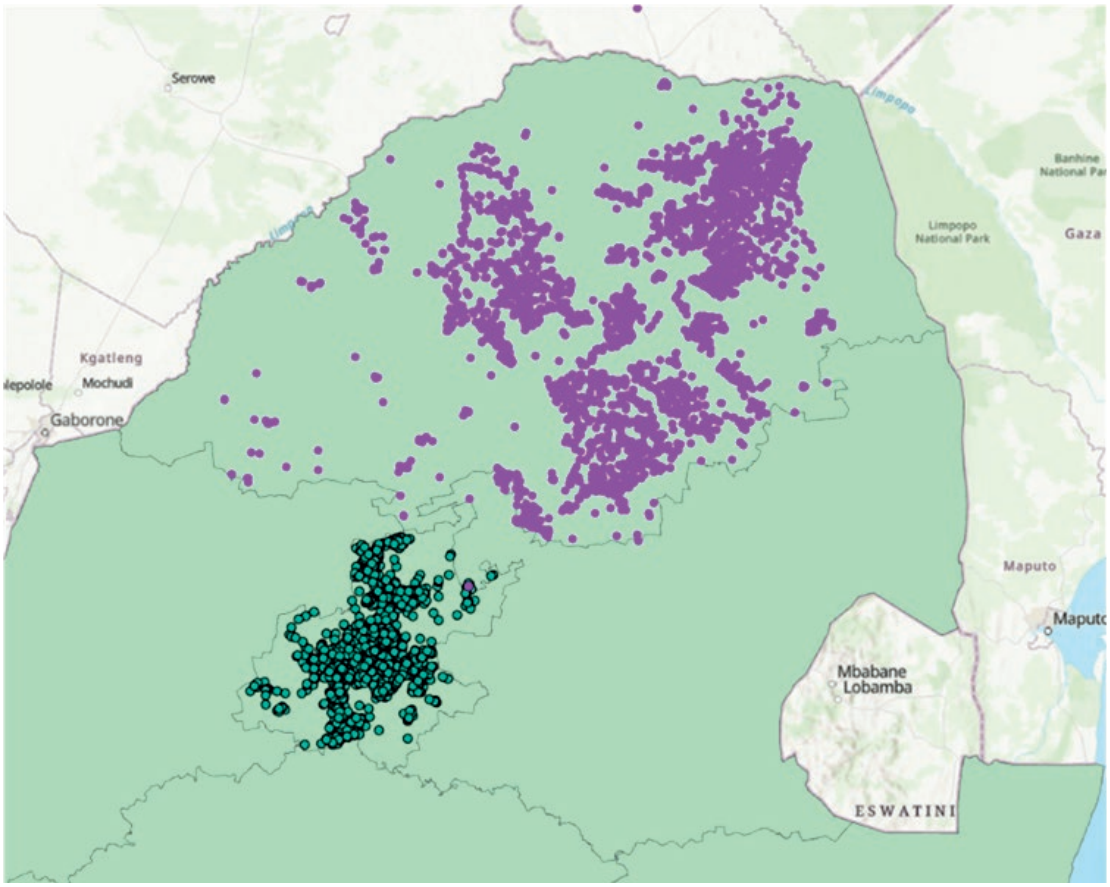
FIGURE 3.5 Outward-mobility of learners from the Eastern to the Western Cape



This section of the report brings to light the dynamic nature of learner across provinces in South Africa, especially under the unique influence of the Covid-19 pandemic. This data provides new insights into mobility patterns and also seem to indicate a smaller influence of the pandemic on such mobility when considering annual movements, which is all that is possible with the Lurits data.

The dot density Map in Figure 3.6 details the movement of learners from Limpopo to Gauteng. Dots in Limpopo indicate schools that learners left in 2020, while dots in Gauteng represent their 2021 destinations.

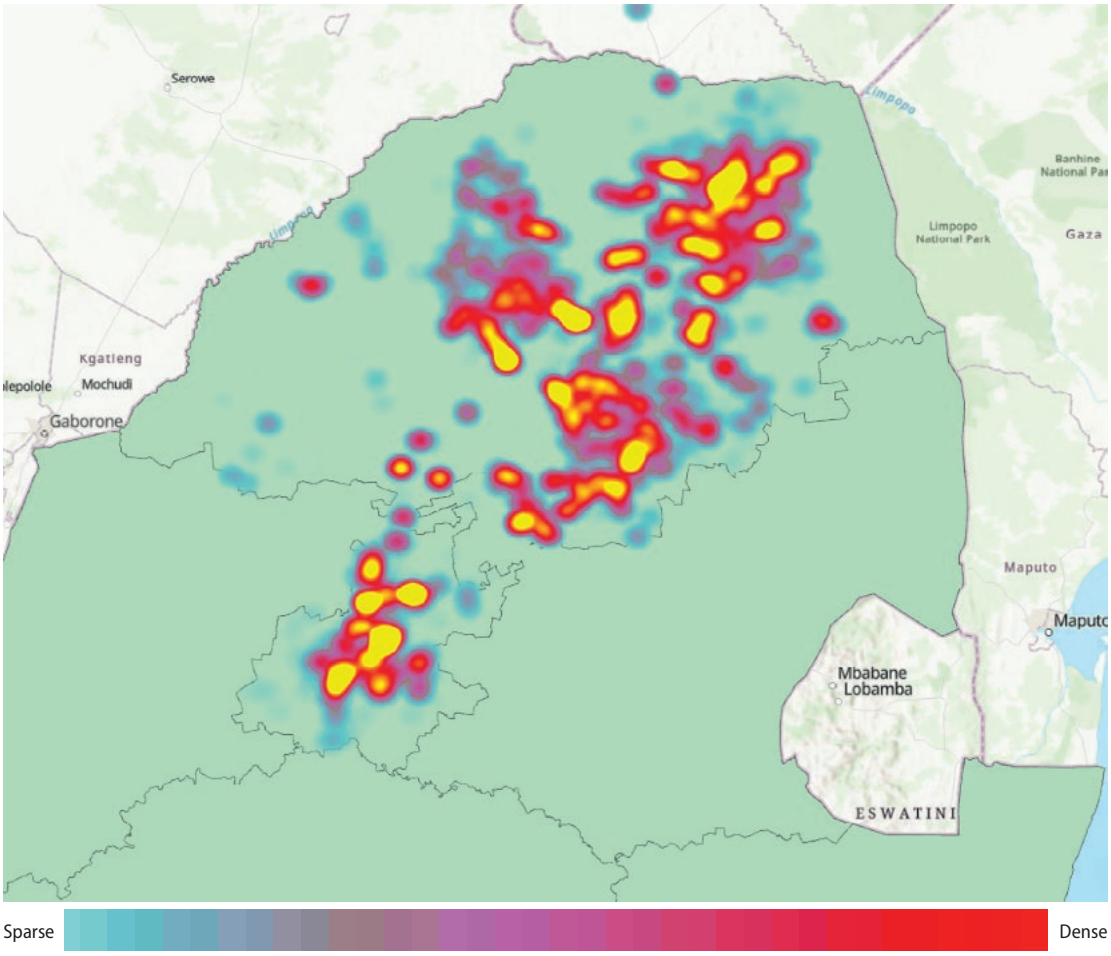
FIGURE 3.6 Dot density map showing schools learners moved from Limpopo in 2020 and the schools they went to in Gauteng in 2021



Source: Based on DDD data

The heat map in Figure 3.7 is a data visualization technique that uses colours to show the extent of change. In Figure 3.7, the yellow shows the high-density areas from where most of the learners came that moved to schools in Gauteng in 2021, while the same colours indicate the numbers of learners migrating to different areas in Gauteng.

FIGURE 3.7 Heat map showing the schools that learners moved from Limpopo in 2020 to schools they went to in Gauteng in 2021



Source: Based on DDD data

3.4 School mobility of Gauteng's 2018 grade 8 cohort

Cohort analysis of grade 8 Gauteng learners of 2018 to 2022 shows that only small numbers of learners switch schools more than once. Although 13% of learners moved once (most often between grades 9 and 10), only about 2% switched more than once, implying that 85% of grade 8 children did not change schools at all at secondary level.

3.5 Mobility between public and private (independent) schools

Anecdotal evidence indicates that financial considerations reduced moves from public to private schools during and after the pandemic. The Lurits data shows shifts between schools for three years. In net terms, there was a small shift of learners to private schools in 2018–19 and a negligible shift in the same direction in 2019–20. In 2020–21, i.e. after the outbreak of the pandemic, a positive shift towards private schools only occurred in the two wealthiest provinces, Gauteng and the Western Cape, while other provinces saw a net movement away from private schools. This therefore supports the contention that the financial crisis caused by the pandemic had an influence on school choice.

TABLE 3.6 Mobility of learners between public and independent (private) schools within province, 2018–2021

	Public to private			Private to public			Net switch to private sector		
	2018–19	2019–20	2020–21	2018–19	2019–20	2020–21	2018–19	2019–20	2020–21
EC	8 351	7 300	7 595	7 784	8 494	9 820	567	–1 194	–2 225
FS	2 041	2 013	1 950	2 165	2 260	2 255	–124	–247	–305
GT	19 604	20 494	22 288	16 713	16 471	20 112	2 891	4 023	2 176
KZN	5 979	5 445	5 380	6 325	6 887	7 550	–346	–1 442	–2 170
LP	5 903	6 971	5 722	5 985	6 629	7 856	–82	342	–2 134
MP	2 961	3 290	2 872	3 281	3 461	4 115	–320	–171	–1 243
NC	607	688	449	465	573	649	142	115	–200
NW	1 815	1 414	1 596	2 374	2 155	2 450	–559	–741	–854
WC	3 534	3 932	5 334	4 160	4 521	5 146	–626	–589	188
SA	**50 795**	**51 547**	**53 186**	**49 252**	**51 451**	**59 953**	**1 543**	**96**	**–6 767**

Source: Calculated from Lurits data

CHAPTER 4
USING SA-SAMS DATA TO UNDERSTAND TEACHING PATTERNS IN EASTERN CAPE, GAUTENG AND LIMPOPO

Summary

SA-SAMS has proven invaluable in building a national dataset of individual learners. These data are a vital part of the overall school funding system. Attention is now turning to the utility of data from other SA-SAMS modules. The current report examines data on what educators in public and independent schools teach, in three provinces and over the years 2015 to 2023. What educators specialise in at school has remained largely unknown, and basic questions such as 'How many mathematics teachers do we have in schools?' have remained unanswerable. This obviously makes planning difficult. The current report focusses on the utility and reliability of the educator data that were made available. For two of the provinces, the data appear close to complete for recent years, while for Gauteng the data are around 60% complete. The internal consistency of the data seems high, for instance with respect to educator identifiers and subject descriptions. Breakdowns of educator specialisation emerging from the SA-SAMS data are consistent with figures from completely different data sources published in a teacher supply and demand report of 2020. One drawback with the SA-SAMS data is that to a high degree educators who are missing in the data tend to be precisely those educators more likely to leave the system soon: older educators, and educators without permanent employment. The data in their current form are thus not suitable for calculating attrition rates for educators with specific specialisations. Such rates are important for refining statistics on what the annual demand for new teachers is.

SA-SAMS has proven invaluable in building a national dataset of individual learners.

4.1 Introduction

For over a decade, the design of the South African School Administration and Management System, SA-SAMS, has matured and its use has expanded. SA-SAMS is a management system, issued by the Department of Basic Education (DBE), used by over 92% of public schools in South Africa, with coverage exceeding 97% in six provinces.[15] Currently, live data reside on computers at the school, while there is a tool allowing schools to submit data to a central provincial warehouse when needed. A process is currently under way to make live data reside on central servers and for schools to access the facility through the Web.

SA-SAMS data from provinces are periodically extracted to a national level through two routes: the DBE works with provinces to collate the data; and the Data Driven District (DDD) initiative[16] brings together provincial data to feed a data-querying portal intended primarily to assist managers in the country's 75 education districts.

SA-SAMS has played a vital role in the building of a national database of individual learners, which is now used to determine education sector weights in National Treasury's equitable share system, which drives the funding of provincial governments.[17] The high stakes attached to the individual learner data means that these data can be considered of a high quality. Several controls make it difficult for schools or provinces to create 'ghost learners' in order to increase funding. The data have been successfully used to generate reports on grade repetition and dropping out.[18]

To some extent data from SA-SAMS other than basic learner records have been used to analyse patterns in the schooling sector. For instance, some analysis of the results of assessments run by schools has occurred.[19] However, this has been limited.

The present analysis is the first focussing on a selection of educator variables in the SA-SAMS data. To some extent, this can be seen as an extension of earlier work making use of the now discontinued teacher questionnaire of the Annual Survey of Schools.[20] It appears the most recent national data drawing from those questionnaires are from 2011.[21] Importantly, the focus of the current report is on the use of SA-SAMS educator data for system-wide monitoring and planning purposes. How *schools* can use SA-SAMS to better manage the school's human resources is a separate matter which the report does not deal with directly. This question is of course important and warrants attention in future.

15 Department of Basic Education, 2019: 134.
16 https://www.home.dbedashboard.co.za
17 Annexure W1 of 2023 Budget Review.
18 Department of Basic Education, 2023.
19 Van der Berg et al, 2019.
20 National Treasury, 2017.
21 Department of Higher Education and Training, 2020: 31.

It should be emphasised that an absence of data on educators has been a stumbling block in the past, with apparently simple questions such as 'How many teachers teach mathematics?' being unanswerable.[22] The Persal payroll system is extensive and generally efficient, but its focus has been fairly strictly limited to information that affects pay.[23] This means that issues such as what an educator teaches are not captured on Persal.

Section 4.2 below describes what educator variables, from the three provinces Eastern Cape, Gauteng and Limpopo, were available for the current analysis.

Section 4.3 discusses the completeness and reliability of the data that became available.

Section 4.4 presents statistics on educators by specialisation in the three provinces. This is what is commonly understood as total demand. This is different from the annual demand for new entrants. The latter does not seem obtainable from the available data, due to incompleteness issues, specifically insufficient replication of records over time to detect attrition patterns (a drawback discussed in section 4.3). Specialisation in terms of phases of the schooling system (section 4.4.1), Foundation Phase language (section 4.4.2) and secondary-level subjects (section 4.4.3) is discussed. The section serves in part to confirm the utility of the SA-SAMS educator data for this type of disaggregation.

Finally, section 4.5 concludes, and reflects on further work that could be done using the SA-SAMS educator data.

4.2 The sub-set of the SA-SAMS data available for the current analysis

Anonymised SA-SAMS educator data for the years 2015 to 2023, for the three provinces, were obtained by RESEP[24] at Stellenbosch University through the DDD initiative. The SA-SAMS data received consisted of 3 921 392 observations, each covering an educator and one of his or her actual teaching specialisations, meaning the grade and subject taught in that year. An educator could appear more than once in a year. Just a few variables out of the total set of educator variables in SA-SAMS were included: year; the official nine-digit school identifier (the 'EMIS number'); who pays the educator's salary; date of birth; grade being taught; and subject being taught. Each educator carried an anonymous unique identifier, which was generated by DDD.

22 This chapter follows the legal descriptions where an 'educator' is anyone employed in terms of the Employment of Educators Act, and a 'teacher' is an educator who is at level 1 according to the Personnel Administration Measures. Nationally, around 78% of educators are teachers. The remaining 22% would have a more senior rank, for instance school-based heads of department, and school principals.

23 National Treasury, 2017: 48.

24 Research on Socio-Economic Policy.

A 2013 description of the SA-SAMS 'Human Resource Information' module[25] confirms that the data made available for the current analysis are just a sub-set, albeit an important one, of the overall data on educators which can be captured within SA-SAMS. The wider set of SA-SAMS variables covers, for instance: qualifications obtained; years of training with respect to specific schools subjects; professional development activities; developmental appraisal; and attendance and leave.

4.3 Data completeness within the available SA-SAMS dataset

A key matter is the integrity of the anonymous educator identifiers in the received data. The level of integrity seems good. There are 195 563 uniquely identified educators in the data of the three provinces, spread across the 3 921 392 observations. The fact that date of birth corresponds perfectly with the educator identifier across years and across provinces suggests that each educator identifier identifies the same educator. A minor problem is that for 790 educator identifiers, the date of birth is missing. There is some repetition of the same educator identifier across different schools in a year, but this is very limited: the number of across-school duplicated educators is on average 58 a year, the highest number being 199 in 2021. Virtually all duplication is within a province. This could be due to educators moving across schools. The pattern of anonymised identifiers is compatible with a situation where the original data had highly accurate 13-digit national identity numbers, and where these were then converted to anonymised values.

Table 4.1 provides further details on the data. Data availability has in general improved over time, both with respect to school and educator coverage. In the case of Eastern Cape, there are no data for 2015. It is striking that the annual increase for educator presence is far greater than that for school presence, indicating more educators in each school are being covered.

25 See https://www.thutong.doe.gov.za/administration/Administration/GeneralInformation/SASAMS/tabid/3346/Default.aspx

TABLE 4.1 Schools and educators in the data

	EC		GP		LP	
Year	Schools	Educ.	Schools	Educ.	Schools	Educ.
2015			820	11 348	1 195	7 394
2016	4 301	24 281	1 494	21 847	3 507	23 093
2017	5 199	34 005	1 801	28 568	3 747	28 818
2018	5 302	39 148	1 921	33 773	3 782	31 921
2019	5 358	43 578	1 855	35 729	3 802	35 583
2020	5 285	45 695	1 608	33 542	3 734	37 727
2021	5 292	51 415	1 764	41 492	3 776	43 224
2022	5 228	57 117	1 980	53 133	3 718	50 491
2023	4 475	52 360	1 261	37 941	3 212	46 236
In any	5 526	62 537	2 725	77 231	3 975	55 819
Annual % increase	0.3	9.6	2.9	11.4	4.3	13.4
% 2022 over Realities	98	91	67	58	96	94
% 'in any' over Realities	103	100	93	84	103	104

The second-last row of Table 4.1 is the 2022 value divided by the corresponding 2022 values published in the DBE's *School Realities* publication, using the statistics for public and independent schools combined in that publication. The year 2022 is of interest as it displays the highest coverage in terms of educators across all three provinces. In the publication, any educator working in a school is included, regardless of who pays them.

In the SA-SAMS data there are only five schools in each of Eastern Cape and Limpopo with school EMIS numbers not found in the official quarter 3 schools master list for public and special schools published online, and one for Gauteng. Independent schools are rather well covered in Eastern Cape and Limpopo, with around four-fifths of independent schools from the master list appearing in the SA-SAMS 2022 data.

In Gauteng, where the use of SA-SAMS is a bit lower, in part due to the relatively strong presence of 'third party' systems, 75% of public schools and 39% of independent schools are covered in the 2022 SA-SAMS data. Surprisingly, better off public ordinary schools in Gauteng are *more* likely to be present in the SA-SAMS data: 82% of quintile 5 schools are covered, around 75% of quintiles 3 and 4 schools, and just 65% of quintiles 1 and 2 schools

(in 2022). This is surprising if it is assumed that third party systems are more likely to be used by more middle class schools and if it is assumed that this is a key reason for SA-SAMS data not being readily available. It should be kept in mind, however, that quintiles 1 and 2 schools account for just 28% of Gauteng's public school enrolment, a lower level than the 46% seen nationally.

The last row of Table 4.1 indicates schools and educators found in *any* year of the SA-SAMS data, divided by 2022 *School Realities* figures. Here the percentage could exceed 100%, for instance due to schools existing in an earlier year but not 2022.

Table 4.2 below illustrates the 2022 educator remuneration patterns seen in the 2022 SA-SAMS data. 'R3' is a category commonly taken to mean not paid by the state nor the school governing body (SGB). These could be educators paid by some NGO, for instance. Virtually all educators in independent schools are *not* paid by the state, as one would expect. SGB-paid educators in quintiles 4 and 5 schools appear to be better represented in the SA-SAMS data than state-paid educators, especially in Eastern Cape and Limpopo.[26] This is perhaps because schools have an incentive to maintain the records of SGB-paid educators as the school must pay their salaries. It is noteworthy that especially in Eastern Cape, SGB-paid educators are relatively common even in quintiles 1 to 3 schools, though officially these schools should not charge fees. This is an observation that has been made previously.[27]

TABLE 4.2 Remuneration patterns and school type in 2022

	EC			GP			LP		
	Ed. count	% SGB	% R3	Ed. count	% SGB	% R3	Ed. count	% SGB	% R3
Public Q1	16 876	8	2	4 965	1	0	16 063	1	1
Public Q2	10 658	7	2	5 440	1	1	18 601	1	1
Public Q3	20 274	6	1	9 952	20	8	9 854	1	1
Public Q4	2 102	30	0	9 023	5	1	754	23	1
Public Q5	3 711	47	0	18 355	28	0	1 943	42	0
Other public	681	10	1	1 104	8	0	426	4	1
Independent	2 822	77	23	4 316	78	20	2 834	76	24
Total	**57 124**	**14**	**3**	**53 155**	**21**	**4**	**50 475**	**7**	**2**

Note: Totals here are marginally higher than those in Table 1 because to a very limited degree the same educator could be found in different schools, as discussed previously.

26 Table 2 implies there are 14 898 SGB-paid educators across the three provinces, while Department of Basic Education (2018: 12) points to a figure of 12 635. It is quite possible that the difference is due to an increase in SGB-paid educators in public schools in recent years.

27 Department of Basic Education, 2018.

An important question is whether the data are sufficient for gauging educator attrition by the specialisation of the educator, something that matters in the planning of teacher training at universities. It soon became apparent that educator attrition in the SA-SAMS data was much lower than other sources would lead one to expect. To gain a clearer sense of what might explain this, through collaboration with the DBE educators from the SA-SAMS data were linked to Persal data, for the year 2018, using the school EMIS number and date of birth. An earlier year was selected as one aim was to gauge attrition over several subsequent years. It was found that 89 631 individuals could be linked across the two datasets. On the SA-SAMS side, only state-paid educators were considered eligible for linking. This meant 35 366, 29 119 and 30 335 eligible educators from Eastern Cape, Gauteng and Limpopo respectively, of whom 92, 95 and 97 percent could be linked, giving the total of 89 631 linked educators. These percentages can be considered satisfactory. Factors that would prevent the percentages from being 100% would include the fact that the two datasets do not reflect exactly the same point in 2018, and multiple educators in a school with the same birthday. The following five conclusions were drawn from the data linking exercise, and specifically 4.3 below.

Around a third of state-paid educators per public school are missing (in 2018). The row 'Persal same schools' reflects all educators in Persal present in the schools from which the linked educators are from. In total, 10 178 schools appear among the linked educators. Just 60% (89 631 over 148 392) of publicly paid educators in the 10 178 schools could be linked, so it can be assumed that a little over a third of publicly paid educators are missing from the SA-SAMS data (this is if one takes into account the abovementioned barriers to linking).

Educators missing from SA-SAMS tend to be older educators. Overall, the average age for linked educators is 45.7, against 46.5 for all educators in Persal. 'Persal all' in the table means all educators in a school, and excludes educators based in, for instance, district offices. Very noteworthy is the fact that educators above age 55 are especially likely to be missing in the SA-SAMS data: the percentage of educators aged over 55% is 16% in the Persal data, whether one considers schools with linked educators or all schools, but just 7% in the SA-SAMS data. On the other hand, educators aged 30 and below comprise similar proportions of educators among linked educators and all Persal educators. In short, the SA-SAMS data reflects a noteworthy under-representation of older state-paid educators.

Except in Eastern Cape, temporary educators are slightly under-represented in the SA-SAMS data extract, relative to Persal. While the percentage of educators who are women does not vary much depending on which slice of the data is examined, the percentage of educators who are permanently employed is slightly higher in the SA-SAMS data (96%) than in the Persal data (94%, counting same schools). This gap is largely driven by missing temporary educators in Limpopo.

Educators in promotion posts are slightly under-represented in the SA-SAMS data. The percentage of educators who are level 1 teachers is slightly higher in the SA-SAMS data than in the Persal data (80% against 78%), meaning manager educators are somewhat more likely to be missing in SA-SAMS.

SA-SAMS appears to have concentrated on educators who are unlikely to leave public employment. What is especially noteworthy is how very unlikely the linked educators are to have left public employment after 2018, up to 2022. Of the 89 631 educators linked using the 2018 data, virtually all remain in Persal in 2022 (the 100% figure seen in the table in the 'Total' row is actually 99.6%, so a tiny number of educators did leave). Some fall out of the SA-SAMS data: 89% of the original 89 631 remain in 2022. This can be attributed to the general incompleteness of the SA-SAMS data and is not a reflection of actual attrition. Attrition is almost zero, according to Persal, for linked educators. The future presence of an educator in Persal was gauged by searching for the educator anywhere in the education system of the *same* province. If educators who are not linked, but who are in schools with some linked educators, are considered, then according to Persal there is an attrition of 19% between 2018 and 2022 (see 81% in the table). Explanations regarding this peculiarity would have to fall into one of two categories. Firstly, the pattern could have to do with how the data were extracted by DDD. Possibly, an attempt was made to exclude educators who left. This could explain the low percentage of educators aged over 55. Secondly, it is possible that the patterns seen have to do with how records are entered, or removed, in SA-SAMS. For instance, the emphasis in recent years could have been on recording permanently employed and relatively young educators, thus educators who are less likely to leave. It is possible that a recently appointed educator attaining permanent status would be especially likely to be recorded in SA-SAMS. What the patterns do *not* suggest is that educators leave, and then records are not removed from SA-SAMS. If this were the case, then linked educators would not display a 100% presence in Persal in 2022. The patterns suggest that even *non*-linked SA-SAMS educators are unlikely to leave. This can be seen in the fact that future presence for linked SA-SAMS educators is very similar to future presence for all SA-SAMS educators – see the row 'SA-SAMS all'.

TABLE 4.3 Patterns seen after linking SA-SAMS to Persal for 2018

		Educ. count	Schools	Educ. per school	Avg. age end 2018	% over 55	% 30 or below
EC	SA-SAMS all	39 148	5 302	7.4	45.2	6	9
	Linked (SA-SAMS)	32 447	4 989	6.5	45.7	5	8
	Linked (Persal)	32 447	4 989	6.5	45.7	5	8
	Persal same schools	51 057	4 989	10.2	46.8	15	8
	Persal all	52 557	5 540	9.5	46.9	15	8
GP	SA-SAMS all	33 773	1 921	17.6	43.4	9	17
	Linked (SA-SAMS)	27 653	1 594	17.3	44.0	8	16
	Linked (Persal)	27 653	1 594	17.3	44.0	8	16
	Persal same schools	47 607	1 594	29.9	44.7	15	16
	Persal all	67 163	2 216	30.3	44.7	16	16
LP	SA-SAMS all	31 921	3 782	8.4	47.1	7	8
	Linked (SA-SAMS)	29 531	3 595	8.2	47.2	7	7
	Linked (Persal)	29 531	3 595	8.2	47.2	7	7
	Persal same schools	49 728	3 595	13.8	48.1	18	8
	Persal all	51 569	3 934	13.1	48.1	18	8
Total	SA-SAMS all	104 842	11 005	9.5	45.2	7	11
	Linked (SA-SAMS)	89 631	10 178	8.8	45.7	7	10
	Linked (Persal)	89 631	10 178	8.8	45.7	7	10
	Persal same schools	148 392	10 178	14.6	46.5	16	11
	Persal all	171 289	11 690	14.7	46.4	16	11

% women	% perm.	% level 1	% present 2019	% present 2020	% present 2021	% present 2022
			98	96	96	96
			98	97	97	96
73	98	77	100	100	100	100
72	98	75	95	92	86	82
72	98	75	95	91	86	82
			78	67	69	73
			77	67	69	74
76	91	81	100	100	100	99
74	89	79	94	90	86	81
74	89	78	94	91	86	81
			98	96	97	95
			98	97	97	95
63	98	84	100	100	100	100
62	94	82	95	91	86	80
62	94	82	95	91	86	80
			92	87	88	88
			92	88	89	89
71	96	80	100	100	100	100
69	94	79	95	91	86	81
70	93	78	95	91	86	81

Grade R teachers, often called 'Grade R practitioners', are a special case as they have different conditions of service compared to all other educators, and currently do not need to hold university degrees, meaning they are not yet a major concern in the university planning system. Table 4.4 provides some details regarding Grade R teaching in 2022. The data suggest that Grade R teachers were included to a fairly large degree. To illustrate, in Limpopo 2 708 were marked as teaching Grade R in 2022, of which 1 828 were marked as teaching only Grade R. However, when full-time equivalent teachers are calculated, a different picture emerges: the total number of teachers teaching Grade R drops from 2 708 to 2 155. The 2 155 value is calculated by counting fractions of teachers. For instance, if a teacher teaches two subjects in Grade R and three subjects in Grade 1, then Grade R benefits from the presence of 40% of that that teacher. This is not an ideal method, but in the absence of data on hours spent teaching, it can be considered a reasonable proxy. By definition, the 1 828 value would not change as teachers who only teach Grade R are counted as 100% of a teacher. The finding, using fractions of teachers, is that 85% of those who teach Grade R, teach only Grade R. An examination of what subjects are taught by those who teach several grades does not reveal striking patterns. Those who teach several grades and those who teach just Grade R tend to teach similar subjects in Grade R. One explanation would be multi-grade classes. If Grade 1 teachers are compared to Grade R teachers, using the fractions approach for both, the percentages in the second-last row emerge. In Limpopo, Grade R teachers are 74% as numerous as Grade 1 teachers. Yet Grade R enrolments in public schools come to 91% of Grade 1 enrolments. This suggests that some Grade R teachers are excluded from the Limpopo data. The percentages for the other two provinces are more compatible with a balanced dataset: in both Eastern Cape and Gauteng Grade R absorbs a higher proportion of teachers than enrolments would suggest, which is believable if the assumption that Grade R classes are a bit smaller than Grade 1 classes holds true.

TABLE 4.4 Grade R teaching in public schools in 2022

	EC	GP	LP
Educators teaching Grade R	4 487	2 925	2 708
Number of above teaching only Grade R	3 092	2 706	1 828
% from above	69	93	68
Educators teaching Grade R (sum of fractions)	3 548	2 755	2 155
% from above	87	98	85
Educators teaching Grade 1 (sum of fractions)	3 960	3 525	2 919
Grade R over Grade 1 teachers (fraction-based)	90	78	74
Grade R over Grade 1 enrolment (%)	76	57	91

Turning to subjects, there are 313 unique subject descriptions in the data. Several carry the prefixes 'TO', 'S' or 'P', which seems to point to a classification of the subject relating to whether a special needs or vocational focus applies. After these prefixes are removed, the total number of unique subject descriptions drops to 231. The prefixes seem relatively unimportant in terms of the aims of the current analysis: an isiZulu teacher in a special school must still specialise in teaching isiZulu when training as a teacher, even if there may be specific pedagogical requirements. Table 4.5 breaks down the 231 subject descriptions by grade and category indicated by the prefix. In the SA-SAMS data, 99.4% of observations have a subject description without a prefix. The most widespread prefix, where these exist, is 'TO', probably standing for technical orientation. These 71 'TO' subjects appear only for grades 6 to 9 teaching. Several of these 'TO' subject descriptions, such as 'Leather work', do not appear in any other of the three categories represented by columns in Table 4.5.

TABLE 4.5 Number of subject values in the raw data (all years)

Grade	General	Prefix 'TO'	Prefix 'S'	Prefix 'P'
R	30 (4.9)		25 (0.02)	14 (0.01)
1	38 (8.1)		25 (0.01)	14 (0.01)
2	37 (7.9)		25 (0.01)	13 (0.01)
3	37 (8.0)		25 (0.01)	14 (0.01)
4	45 (8.4)		25 (0.01)	14 (0.01)
5	44 (8.2)		25 (0.01)	14 (0.01)
6	43 (8.0)	70 (0.11)	25 (0.01)	14 (0.01)
7	52 (11.1)	70 (0.10)	25 (0.01)	14 (0.01)
8	60 (6.9)	60 (0.07)	46 (0.02)	14 (0.01)
9	62 (6.7)	54 (0.06)	46 (0.02)	14 (0.01)
10	112 (7.4)		46 (0.02)	14 (0.01)
11	117 (7.0)		46 (0.02)	14 (0.01)
12	143 (6.6)		46 (0.02)	14 (0.01)
Total	**182 (99.4)**	**71 (0.34)**	**46 (0.20)**	**14 (0.09)**

Note: Values in brackets in this and the next table are the percentage of all the 3 921 392 observations in the dataset.

It seemed logical, in terms of the need to analyse educator demand, to reduce duplication among the 231 subject descriptions. For instance, home, first additional and second additional languages could be described simply as the language concerned, for instance 'Xitsonga'. There were many very specialised music descriptions which were all renamed 'Music'. This reduced the number of subject descriptions to 108. The breakdown of the 108 is shown in Table 4.6 below, by grade and the language status of the subject. There are 12 official language descriptions as one is 'South African Sign Language'. Foreign languages account for a tiny percentage of all observations in the data, just 0.2%. Two-thirds of observations are accounted for by the teaching of non-language subjects, and a further one-third by the teaching of an official language.

TABLE 4.6 Number of subject values after normalisation (all years)

Grade	Non-language	Official language	Non-official language
R	5 (3.2)	12 (1.7)	
1	3 (4.0)	12 (4.1)	1 (0.003)
2	3 (3.9)	12 (4.0)	
3	3 (4.0)	12 (4.1)	
4	6 (5.5)	12 (3.0)	3 (0.002)
5	6 (5.3)	12 (2.9)	3 (0.002)
6	41 (5.3)	12 (2.8)	3 (0.002)
7	41 (8.6)	12 (2.7)	3 (0.002)
8	45 (5.2)	12 (1.8)	3 (0.002)
9	46 (5.0)	12 (1.8)	3 (0.001)
10	41 (5.6)	12 (1.8)	16 (0.054)
11	41 (5.3)	12 (1.7)	16 (0.052)
12	41 (5.0)	12 (1.6)	23 (0.058)
Total	**73 (65.9)**	**12 (33.9)**	**23 (0.179)**

The 108 subject descriptions covered in Table 4.6 are descriptions found across the nine years of the data. Focussing only on the most recent year, 2023, yields 95 subjects, largely because there are fewer foreign languages present in just that year.

There are large differences in the degree of presence of each of the 108 subjects. For instance, the largest 19 subjects, in terms of observations in the data, account for 90% of all observations. Just 35 of the 108 subjects account for 99% of all observations.

4.4 Using the data to gauge total specialisation-specific demand

4.4.1 Phase-specific demand

A relatively straightforward and important point of departure is to gauge the spread of educators across the levels that educators specialise in. Levels are based on grades, and this is the focus of Table 4.7 below. Two similar tables appear in the 2020 DHET teacher supply and demand report, one based on 2011 Annual Survey of Schools (ASS) teacher data, and the other on 2017 sample-based data.[28] The 2011 teacher data, which in many ways is the precursor to the SA-SAMS educator data, pointed to 45% of grades 1 to 12 teachers nationally doing some teaching at the grades 8 to 12 secondary level.[29] The figure becomes 41% using the 2022 SA-SAMS data across the three provinces, and 40% if the analysis is limited to publicly paid educators. In calculating these statistics the 'non-major combinations' referred to in Table 4.7 were considered. The relatively low 41% figure is in part explained by the fact that the secondary level is relatively small in the three provinces: in these three, grades 8 to 12 enrolments came to 64% of grades 1 to 7 enrolments in 2022, against 67% for the other six provinces. (Comparison against the earlier analysis of the 2017 data is made difficult by the fact that the earlier analysis did not weight schools, yet secondary schools were over-sampled relative to primary schools.)

TABLE 4.7 Grade combinations of educators in 2022

Grades													Cumulative % of educators		
R	1	2	3	4	5	6	7	8	9	10	11	12	EC	GP	LP
⊗													6	6	4
	⊗												11	13	9
		⊗											16	20	15
			⊗										22	27	21
				⊗									24	31	23
				⊗			⊗						25	32	25
				⊗		⊗	⊗						25	33	26
				⊗	⊗								27	35	28
				⊗	⊗		⊗						28	37	30
				⊗	⊗	⊗							33	38	32
				⊗	⊗	⊗	⊗						41	40	38

28 Tables 12 and 13 of Department of Higher Education and Training (2020).
29 This 45% can be calculated from Table 12 in Department of Higher Education and Training (2020).

Grades													Cumulative % of educators		
R	1	2	3	4	5	6	7	8	9	10	11	12	EC	GP	LP
					⊗								43	43	40
					⊗		⊗						44	45	42
					⊗	⊗							45	47	43
					⊗	⊗	⊗						47	49	46
						⊗							48	51	47
						⊗	⊗						51	55	50
							⊗						54	59	53
							⊗	⊗	⊗				56	60	54
Non-major combinations starting in grades R to 7													68	65	60
								⊗		⊗	⊗	⊗	69	66	62
								⊗	⊗				72	69	64
								⊗	⊗	⊗			73	71	66
								⊗	⊗	⊗	⊗		74	73	67
								⊗	⊗	⊗	⊗	⊗	79	77	74
									⊗	⊗	⊗	⊗	80	78	76
										⊗	⊗		82	80	77
										⊗	⊗	⊗	88	85	82
											⊗	⊗	90	88	84
												⊗	91	89	85
Non-major combinations starting in grades 8 to 12													100	100	100

Source: SA-SAMS

Note: Values in the last three columns are cumulative percentage values, and sorting of the initial 13 columns, representing grades R to 12, follows the patterns of the 'X' values, with 'X' values for lower grades appearing first. To illustrate, the table indicates that 56% of educators in Eastern Cape were teaching one of the major 19 combinations shown in the first 19 rows, involving some teaching in grades R to 7. Any combination consisting of at least 1.0% of educators was considered 'major' – here the average percentage across the three provinces, unweighted, occurred. In Eastern Cape, a further 12% of educators (68% minus 56%) were teaching some 'non-major' combination of grades, where the combination had to include at least one grade at the primary level. Within this 12% are some 502 combinations found across the three provinces. Within the non-major combinations confined to the secondary level, accounting for 9% of educators in Eastern Cape, are 21 combinations. The total educators considered for the table are 57,135, 53,156 and 50 500 for EC, GP and LP.

Table 4.8 provides a similar analysis, except here the phases of the schooling system are considered. Around two-thirds of educators in Eastern Cape and Gauteng were confined to just one phase, while this was a lower 51% in Limpopo.

TABLE 4.8 Phase combinations of educators in 2022

Foundation	Intermediate	Senior	FET	EC	GP	LP
✕				27	28	23
	✕			14	17	11
	✕	✕		17	13	20
		✕		10	9	6
		✕	✕	17	19	28
			✕	12	12	11
			Non-major combinations	3	1	2
				100	100	100

Note: Here and in similar tables below, the sorting of rows is in descending order of the statistic across all three provinces.

As explained in the 2020 DHET report,[30] the relevant 2015 policy requires educators to be ready to teach any grade in one of three levels: the grades R to 3 Foundation Phase; the primary grades 4 to 7; or the secondary grades 8 to 12. Demand and supply should thus be conceptualised in terms of these three levels.

The following table uses the SA-SAMS data to arrive at numbers of educators at each of the major three levels. One adjustment is made. Because Grade R teacher are not required to have a university training, they appear in a row of their own. There are 160 741 educators in the data.[31] The higher total of 171 202 in the first column emerges as some educators are counted more than once, if they teach at more than one level. The 20 181 total also includes duplicates, as this column considers educators teaching across multiple levels. These educators tend each to have more grade-subject records. The second-last column uses fractions of educators, discussed in section 4.3. The last column of Table 4.9 provides a percentage breakdown which is not that different from the percentage breakdown of the first column, where educators were duplicated across rows. These breakdowns consider only grades 1 to 12, the grades which are currently of particular interest to teacher trainers at universities.

30 Department of Higher Education and Training, 2020: 25.

31 This total is 50 teachers lower than what Table 1 indicates. This is because that table has some teachers appearing in more than one province.

TABLE 4.9 Educators per level in the three provinces combined in 2022

	Educators teaching at this level (multiple counting of same)	% break-down	Average grade-subject records of these educators	Educators not only teaching at this level	Average grade-subject records of these educators	Educators per level using fractions approach (no duplication)	% break-down
Grade R	11 101		5.1	2 677	11.8	9 323	
Grades 1 to 3	36 081	22.5	5.0	4 329	10.8	34 092	22.5
Grades 4 to 7	61 217	38.2	4.4	7 959	8.1	56 815	37.5
Grades 8 to 12	62 803	39.2	4.0	5 216	6.6	60 512	40.0
Total	**171 202**	**100.0**	**4.1**	**20 181**	**8.2**	**160 741**	**100.0**

Table 4.10 provides educators per level, using the fractions approach, for each province separately. The sum across the three province is still 160 741, as in the case of the very small number of educators appearing in different provinces, educators are split across provinces as fractions, in line with each educator's number of records.

TABLE 4.10 Educators per level in each of the three provinces in 2022

	Educators (using fractions)			% breakdown		
	EC	GP	LP	EC	GP	LP
Grade R	3 767	3 145	2 411	7	6	5
Grades 1 to 3	12 452	11 999	9 641	22	23	19
Grades 4 to 7	20 677	18 560	17 578	36	35	35
Grades 8 to 12	20 222	19 431	20 859	35	37	41
Total	**57 117**	**53 134**	**50 490**	**100**	**100**	**100**

Table 4.11 below compares the breakdown of educators by level seen in Table 4.9 against that appearing in the 2020 supply and demand report of the Department of Higher Education and Training (DHET). The differences across the last two columns of Table 4.11 are not worrying if one considers they draw data from different years and on a different set of provinces (all nine against just three), and the missing data problem in SA-SAMS discussed above. Official enrolment reports make it clear that the proportion of learners at the secondary level has been rising, which could in part explain the higher percentage of educators at the secondary level in the more recent SA-SAMS data.

TABLE 4.11 Comparison against 2020 DHET report

	Total demand 2023 DHET report		% from Table 4.9
	Educator count	%	
Grades 1 to 3	103 842	27	22.5
Grades 4 to 7	137 221	35	37.5
Grades 8 to 12	150 122	38	40.0
Total	**391 184**	**100**	**100.0**

Source: The DHET values are the 2023 values behind Figure 29 in Department of Higher Education and Training (2020: 59).

4.4.2 Language-specific demand in the Foundation Phase

Table 4.12 draws from the records of 35 165 educators teaching some official South African language in grades 1 to 3 across the three provinces. Here full educators, and not fractions of educators, are used. Nineteen of the educators, split roughly equally across Eastern Cape and Limpopo (none in Gauteng), taught South African Sign Language. This language is not considered in Table 4.12. The table reveals the patterns one might expect from the data. For instance, in Limpopo 84% of Foundation Phase educators who teach a language, say they teach one African language as a home language (HL), and then English as an additional language. For the purposes of this analysis, 'first additional language' and 'second additional language' were simply classified as 'additional'. Second additional language has a tiny presence: of all the 'additional' records in the data analysed here, only 0.8% are 'second additional'. The third row in the table is Afrikaans HL combined with English additional. The fourth is English HL combined with an African language as additional. This is relatively common in Gauteng, with 9% of educators analysed having this combination. This could reflect success in getting more children without an African language as a home language to take this language, in line with the 2013 Incremental Introduction of African Languages (IIAL) policy. But it could also reflect independent schools where parents with an African home language have opted for English home language to be used in the initial grades. There are very few educators teaching more than one African language – this would be reflected by values greater than 1 in the first two columns.

TABLE 4.12 Language combinations of Foundation Phase educators 2022

Languages (only first two columns can be greater than 1)						Cumulative % of educators		
African HL	African add.	Eng HL	Eng add.	A'kaans HL	A'kaans add.	EC	GP	LP
1			1			71	35	84
		1			1	80	68	87
			1	1		87	78	89
	1	1				89	87	90
		1				91	93	93
1						93	93	95
			1			95	94	98
	1					97	95	98
2			1			98	96	99
					1	98	97	99
	1	1			1	99	98	99
43 smaller combinations						100	100	100

For Table 4.13 below, any Foundation Phase educator reportedly teaching a *home language* was considered. The home language is critical as according to the curriculum this should be the language of learning and teaching (LOLT) used by the educator to teach the non-language subjects mathematics and life skills. Extracting just home language teaching meant 33 945 of the abovementioned 35 165 educators were available for the analysis. The 33 945 dropped to 32 539 when the fractions approach was employed to deal with educators teaching not just in the Foundation Phase. There were cases where educators taught more than one home language in the Foundation Phase. This duplication produced 547 'duplicate' records. By far most of this phenomenon consisted of educators with two home languages taught. In cases such as this, the count of the educator (using the fractions approach) was split equally across the languages taught. Table 4.13, like Table 4.11 below, points to the specialisations in SA-SAMS being credible. The largest difference across the two sources, SA-SAMS and the rather old 2013 EMIS[32] data used for the 2020 DHET report, relates to English in Gauteng: 52% of the total demand in SA-SAMS, against 43% in the DHET report. This would partly be explained by the fact that the DHET report focusses just on public schools, while SA-SAMS includes independent schools.

32 Education Management Information Systems.

TABLE 4.13 Comparison against Foundation Phase languages in DHET report

	SA-SAMS			2020 DHET report		
	EC	GP	LP	EC	GP	LP
Afrikaans	7	11	2	6	8	1
English	16	52	8	13	43	6
isiNdebele		0	1	0	0	0
isiXhosa	75	2	0	79	3	0
isiZulu		14	1	0	18	1
Sepedi		6	56	0	8	58
Sesotho	2	6	0	2	8	0
Setswana		7	1	0	9	1
siSwati				0	0	0
Tshivenda		1	16	0	0	16
Xitsonga		2	16	0	2	17
Total	**100**	**100**	**100**	**100**	**100**	**100**
Educators	**11 793**	**11 496**	**9 250**			

Source: DHET report figures are from Table 2 of Department of Higher Education and Training (2020: 12). The DHET report figures refer specifically to LOLT.

Note: Here and in similar tables that follow, for SA-SAMS statistics zero is a low non-zero value below 0.5%, while missing is truly zero.

4.4.3 Subject-specific demand at the secondary level

As shown in Table 4.9 above, in 2022 there were 62 803 educators in the three-province SA-SAMS dataset teaching grades 8 to 12, which declines to 60 512 if fractions of educators also teaching outside this grade range are used. These educators taught 91 subjects, out of the 108 normalised subjects discussed in relation to earlier Table 4.6. These 91 consisted of 67 non-language subjects, 11 official languages, and 13 non-official languages (here the breakdowns of Table 4.6 are used).

The following two tables illustrate that teaching one or two subjects in either the grades 8 to 9 range (Table 4.14), or the grades 10 to 12 range (Table 4.15), is the norm. Strikingly, the first seven of the eight rows in each table are identical with respect to type of combination. Patterns are clearly extremely similar across the two grade ranges. As for Table 4.12 above, combinations are sorted from most to least common, using the average across the three provinces. Most common is teaching one non-language subject, and nothing else. This is true for around half of educators. Second-most common is teaching one official language,

and nothing else. The first two combinations on their own account for around two-thirds of educators. A column for teaching non-official languages would have been included in the tables if the additional existing combinations had been specified – for instance, there are 34 additional combinations for Table 4.14, beyond the eight shown, with these eight accounting for around 99% of all educators.

TABLE 4.14 Subject counts of grades 8 to 9 educators in 2022

		Cumulative % of educators		
Non-language	**Official language**	**EC**	**GP**	**LP**
1		45	53	45
	1	64	77	64
2		80	88	80
1	1	91	96	91
3		94	97	94
2	1	97	98	97
	2	98	99	98
3	1	98	99	98
34 smaller combinations		**100**	**100**	**100**

Note: In this table and the following two, the fractions approach was not employed. Moreover, an educator teaching English to two grades and history to one grade would be counted as one English and one history teacher.

TABLE 4.15 Subject counts of grades 10 to 12 educators in 2022

		Cumulative % of educators		
Non-language	**Official language**	**EC**	**GP**	**LP**
1		49	53	46
	1	69	77	65
2		87	90	83
1	1	95	96	94
3		98	98	97
2	1	99	98	98
	2	99	99	99
1	2	100	99	99
46 smaller combinations		**100**	**100**	**100**

Table 4.16 below extracts from the data the actual subjects taught. To illustrate, the most common combination is teaching only English. This accounts for 8% of grades 8 to 12 educators in Eastern Cape. As explained earlier, the subject normalisation process collapsed English home language and English first additional language into just English. Universities appear not to make the distinction in their training programmes, judging from the course descriptions of a few universities. The second-most common combination is teaching just mathematics. Here mathematics is distinct from the subject mathematical literacy. Universities tend to make the distinction when training. Though Table 4.16 details only 50 combinations, out of a total of 3 684, the largest 50 account for around up to two-thirds of all educators (see Gauteng).

TABLE 4.16 Subject combinations of grades 8 to 12 educators 2022

		Cumulative % of educators		
Non-language	Official language	EC	GP	LP
	English	8	10	7
Mathem		13	16	11
	Afrikaans	15	22	12
	IsiXhosa	21	22	12
Life O		23	25	13
Mat Li + Mathem		24	28	15
Life O	English	26	30	17
Mat Li		28	32	18
Life S		30	34	20
Life S + Natura		31	36	21
	Sepedi	31	37	25
Geogra		32	38	26
Geogra + Social		33	40	28
Physic		35	41	29
Natura + Physic		35	42	30
Mathem + Physic		36	43	32
Histor		37	44	33
Accoun + Eco Ma		38	45	34
Busine		39	47	34
Histor + Social		40	48	35

		Cumulative % of educators		
Non-language	Official language	EC	GP	LP
Social		41	49	35
Busine + Eco Ma		42	50	36
Mathem + Natura		43	51	36
Natura		44	51	37
Touris		45	53	37
Techno		46	53	37
Agricu		47	54	38
Mathem + Techno		47	54	39
Eco Ma		48	54	39
Eco Ma + Econom		49	55	40
Creati		50	56	40
Social	English	50	56	41
	IsiZulu	50	58	41
Creati	English	51	59	41
Life O	IsiXhosa	53	59	41
Accoun		53	60	42
Econom		54	60	42
Histor	English	55	61	43
Life O	Sepedi	55	61	44
Eco Ma + Life O		55	61	44
Geogra	English	56	61	45
	Tshivenda	56	61	46
Creati + Life O		56	62	47
	Xitsonga	56	62	48
Life S + Mathem		56	62	48
	Setswana	56	63	49
Mathem + Natura + Physic		57	64	49
Comp A		57	64	49
Life O + Social		57	65	50

		Cumulative % of educators		
Non-language	Official language	EC	GP	LP
Life O + Mathem		58	65	50
3 634 smaller combinations		100	100	100

Note: Subject names are the first six letters of the official subject name, except in a few cases where there were ambiguities: 'Mathem' is mathematics while 'Mat Li' is mathematical literacy; 'Econom' is economics while 'Eco Ma' is economic management sciences.

Table 4.16 offers very preliminary insights into how subject teaching is organised in the secondary grades. There is much scope for further analysis of the available SA-SAMS data relating to, for instance, how subjects taught changes from one year to the next, the relationship between subjects taught and school size, differences in patterns across grades and across socio-economic school quintiles, and so on.

Table 4.17 and Table 4.18 below compare total subject-specific demand indicated by the three SA-SAMS provincial datasets, to *national* statistics published in the DHET report. Of the 91 abovementioned subjects, 68 are reflected across the two tables, these being the subjects emerging when subjects taught by just one educator across all three provinces had been excluded.[33] For this analysis fractions of educators were considered. An educator also teaching outside the grades 8 to 12 range became a fraction of an educator, depending on how many of the educator's observations *were* about grades 8 to 12 teaching. In addition, within the grades 8 to 12 range, fractions were employed. An educator teaching just grades 8 to 12, and teaching English in two grades and history in one grade was counted as two-thirds of an English teacher and one-third of a history teacher. There is an implied assumption that the educator spends two-thirds of her time on English, which may not be entirely accurate. However, the data did not permit any other split.

Even subjects listed below are in many cases taught by very few full-time equivalent educators. To illustrate, only 45 of the 91 subjects were taught by more than ten full-time equivalent educators across the three provinces.

The two tables point to a distribution of subject-specific total demand one would expect given the 2020 DHET report. In Table 4.17 the total demand for English teachers is a bit lower in SA-SAMS than in the DHET report, but the roughly two percentage point difference is easily explained by, for instance, the exclusion of provinces such as KwaZulu-Natal and Western Cape. Statistics for the other ten languages are especially sensitive to province. In Table 4.18, the relatively low DHET report values for the grades 8 to 9 subjects economic management sciences and social sciences would largely be explained by the approach taken for the DHET report: it was assumed that half of the teaching of, say, social sciences would be done by teachers trained to teach history or geography in grades 10 to 12,[34] an

33 The exception to this rule is that siSwati appears in Table 17, to produce the full set of 11 languages, though this language was not taught by any teachers in the three provinces in 2022, according to the data.

34 Department of Higher Education and Training, 2020: 55.

arrangement which is promoted by the relevant teacher minimum standards policy. In line with this, the relatively high value in the DHET report for life sciences would in part be explained by the assumption that life sciences teachers would to a large degree teach natural sciences in grades 8 to 9 (even if natural sciences often appears as a separate training area in the university prospectuses). The relatively high value for mathematics in the DHET report may be explained by the assumption used for the report that mathematics classes are significantly smaller than other classes, in line with the policy that distributes teaching posts to schools. This may not hold true in the actual timetabling practices of schools. This could probably be examined further using the SA-SAMS data.

TABLE 4.17 Comparison against secondary subjects in DHET report (Part 1)

	SA-SAMS 2022			2020 DHET report
	EC	GP	LP	
Afrikaans	2.14	6.56	0.74	3.5
English	12.78	13.76	13.19	15.6
isiNdebele	0.00	0.01	0.14	0.1
isiXhosa	10.60	0.61	0.01	2.3
isiZulu	0.01	3.05	0.12	4.4
Sepedi	0.00	1.37	7.76	2.0
Sesotho	0.40	1.14	0.00	0.8
Setswana	0.00	1.77	0.18	1.2
siSwati	0.00	0.00	0.00	0.5
Tshivenda	0.00	0.15	2.27	0.5
Xitsonga	0.00	0.38	2.27	0.7
South African Sign Language	0.02	0.01	0.01	
Arabic	0.01	0.05	0.00	
French	0.01	0.09	0.00	
German	0.00	0.03	0.00	
Gujarati	0.00	0.02	0.00	
Hebrew	0.00	0.01	0.00	
Urdu	0.00	0.02	0.00	

Source: DHET report values are from Department of Higher Education and Training (2020: 17–18). The 'Percentage of all candidates in 2017' column used from that source.

TABLE 4.18 Comparison against secondary subjects in DHET report (Part 2)

	SA-SAMS 2022			2020 DHET report
	EC	GP	LP	
Accounting	2.28	2.40	2.17	2.1
Agriculture	2.72	0.23	3.29	0.1
Beauty	0.01	0.03	0.00	
Bricklaying and Plastering	0.00	0.01	0.00	
Business Studies	2.94	3.54	2.28	3.9
Civil Technology	0.26	0.24	0.19	0.4
Coding and Robotics	0.00	0.01	0.00	
Computer Applications Technology	0.69	1.44	0.25	0.9
Computyping	0.00	0.01	0.00	
Consumer Studies	0.66	0.99	0.26	1.1
Creative Arts	3.47	2.92	3.21	1.3
Dance Studies	0.01	0.05	0.00	0.0
Design	0.03	0.13	0.00	0.0
Digital Technology	0.03	0.01	0.00	
Dramatic Arts	0.08	0.26	0.01	0.2
Early Childhood Development	0.00	0.01	0.00	
Economic Management Sciences	3.71	3.14	3.29	1.3
Economics	2.11	1.87	2.41	2.6
Electrical Technology	0.22	0.24	0.24	0.2
Engineering Graphics and Design	0.43	1.03	0.36	1.3
Geography	3.66	3.75	5.18	5.5
History	3.11	2.97	2.15	2.8
Hospitality Studies	0.12	0.31	0.06	0.2
Industrial Sewing	0.00	0.00	0.00	
Information Technology	0.08	0.35	0.06	0.1
Life Orientation	8.92	8.46	8.55	6.9
Life Sciences	4.88	4.15	5.81	7.2

	SA-SAMS 2022			2020 DHET report
	EC	GP	LP	
Maintenance	0.01	0.02	0.00	
Maritime Economics	0.01	0.00	0.00	0.0
Mathematical Literacy	4.30	5.08	5.01	6.6
Mathematics	10.41	9.98	10.54	8.2
Mechanical Technology	0.26	0.32	0.20	0.3
Motor Mechanics	0.01	0.02	0.01	
Music	0.32	0.20	0.01	0.2
Natural Sciences	4.14	3.75	4.00	2.0
Nautical Science	0.01	0.01	0.00	0.0
Office Administration	0.02	0.01	0.00	
Panel Beating	0.02	0.01	0.00	
Physical Sciences	3.64	3.20	4.77	4.3
Religion Studies	0.06	0.14	0.01	0.1
Sheet Metal Work	0.00	0.01	0.00	
Social Sciences	4.34	3.80	4.29	2.0
Sport and Exercise Science	0.00	0.01	0.00	0.0
Technical Science	0.26	0.23	0.21	
Technology	3.49	2.71	3.20	1.3
Tourism	2.19	2.49	1.20	2.4
Visual Arts	0.07	0.38	0.04	0.3
Welding	0.02	0.02	0.01	
Wholesale and Retail	0.00	0.01	0.00	
Woodworking and Timber	0.01	0.03	0.01	
Total (%)	100	100	100	
Total fractions of educators	**20,218**	**19,417**	**20,854**	

4.5 Conclusion

This report provides initial insights into how SA-SAMS educator data could be used to fill knowledge gaps relating to educators, gaps which adversely affect the planning of teachers. There is clearly scope for further analysis, even of the somewhat limited sub-set of SA-SAMS educator data which became available for the current report. In particular, some preliminary analysis of this data, not reported on above, suggests that important patterns relating to subject teaching versus all-class teaching at the primary level are confirmed by the data. For instance, historically disadvantaged schools appear more inclined not to do all-class teaching, and to have good mathematics teachers concentrate on just this subject across several grades. Historically advantaged schools, on the other hand, which often enjoy the benefit of better trained teachers, can allow all teachers to teach mathematics, which facilitates all-class teaching, where a teacher teaches a class for every subject.

This report has generated many new insights because of the rich data that it builds on.

CHAPTER 5
CONCLUSION AND LOOKING AHEAD

This report has generated many new insights because of the rich data that it builds on. It is important to reflect on the critical insights gleaned from this analysis of the South African education system in the aftermath of the Covid-19 pandemic. The data-driven approach, utilising the SA-SAMS based DDD and LURITS datasets as well as matric examination data provides an unprecedented window into the dynamics of learner flows, assessment, educator specialisation and the broader implications of these factors on educational policy and planning.

Some key takeaways and implications are the following:

1. **Adapting to the new normal:**
   The pandemic has undeniably transformed the educational landscape, highlighting the need for resilience and adaptability in both teaching and learning methodologies. The decreased repetition rates, adjusted assessment strategies and the shift in learner mobility patterns underscore the system's capacity to respond to unprecedented challenges.

2. **The role of teachers:**
   Data from SA-SAMS has illuminated the specialities and distribution of educators across provinces. This information is pivotal for addressing current and future teacher shortages and for ensuring that educational planning is both strategic and responsive to the evolving needs of the system.

3. **Harnessing data for informed decision-making:**
   The consistent use and analysis of educational data have proven to be indispensable tools for policy-making and educational interventions. High-quality administrative data can guide strategic decisions, ensuring that resources are allocated efficiently and effectively.

Looking to the future, several key areas warrant continued attention and action:

1. **Enhancing data quality and accessibility:**
   While significant strides have been made in data collection and analysis, efforts must continue to improve the completeness, accuracy, and accessibility of educational data. This will ensure accurate and relevant data-driven insights.

2. **Addressing educational inequities:**
   The pandemic has highlighted existing inequities within the education system. It is crucial to continue identifying and addressing these disparities to ensure equitable access to quality education for all learners.

3. **Preparing for the future of education:**
   In the evolving educational landscape, there is a need to anticipate and prepare for future challenges and opportunities. This includes embracing technological advancements, innovating teaching methodologies, and ensuring that the education system remains flexible and resilient.

4. **Strengthening educator support and development:**
   Continued investment in the development and support of educators as the backbone of the educational system is essential. This includes addressing challenges highlighted in the report, such as educator attrition and specialization needs.

In conclusion, this report provides a snapshot of the current state of education in South Africa but also lays the groundwork for future inquiries and strategic interventions. Leveraging data-driven insights and fostering collaborative data-gathering and analysis efforts among stakeholders in education are essential elements in building an education system that is robust, equitable and responsive to the needs of all its participants.

References

Ardington, C., Wills, G., & Kotze, J. (2021). COVID-19 learning losses: Early grade reading in South Africa. *International Journal of Educational Development*, 86 (August), 102480. *https://doi.org/10.1016/j.ijedudev.2021.102480*

Branson, N. & Lam, D. (2009). *Education: Analysis of the NIDS Wave 1 Dataset*. Discussion Paper 3. July. National Income Dynamics Study *(NIDS)*.

Brophy, J. (2006). Grade repetition. *International Institute for Education Planning (IIEP), International Academy of Education (IAE)*. UNESCO.

Burkam, D., Lee, V. E., & Dwyer, J. (2009). *School mobility in the early elementary grades: Frequency and impact*. Prepared for the Workshop on the impact of mobility and change on the lives of young children, schools, and neighborhoods. June 29–30, 2009.

Department of Basic Education. (2013). National policy pertaining to the programme and promotion requirements of the National Curriculum Statement Grades R–12. *Government Gazette*, 74.

Department of Basic Education (2018). *Per learner spending inequities in 2016*. Pretoria.

Department of Basic Education (2019). 2017 School Monitoring Survey: Quantitative survey main report. Pretoria.

Department of Basic Education. (2023). Grade promotion, repetition and dropping out 2018 to 2021. DBE: Pretoria.

Department of Higher Education and Training (2020). School teacher supply and demand in South Africa in 2019 and beyond. Pretoria.

Dixon, S. (2018). Student mobility across schools and its links to under-achievement. Working Paper 18/01. Published April 18, 2018.

Eddy, L. (2011). *The effect of student mobility on student achievement*. University of Kentucky Doctoral Dissertations, 177. *https://uknowledge.uky.edu/gradschool_diss/177*

Goldhaber, D., Koedel, C., Özek, U., & Parsons, E. (2021). *Using longitudinal student mobility across schools and districts to identify at-risk student*s. CALDER Policy Brief No. 25–0421. Arlington, Virginia: National Center for Analysis of Longitudinal Data in Education Research.

Gustafsson, M. (206). *Understanding trends in high-level achievement in Grade 12 mathematics and physical science*. Resep: Stellenbosch. *https://resep.sun.ac.za/wp-content/uploads/2016/06/Correcting-racial-imbalances-2016-01-25.pdf*

Hoadley, U. (2020). Schools in the time of COVID-19: Impacts of the Pandemic on Curriculum. *Resep Working Paper*. Resep: Stellenbosch.

National Treasury (2017). Towards better generation and use of data within the basic education sector. Pretoria.

Spaull, N. (2013). Poverty & privilege: Primary school inequality in South Africa. *International Journal of Educational Development* 33(5), 436–447. *https://doi.org/10.1016/j.ijedudev.2012.09.009*

Van der Berg, S., Hoadley, U., Galant, J., van Wyk, C., & Bohmer, B. (2022a). *Learning Losses from COVID-19 in the Western Cape. Evidence from Systemic Tests*. Resep: Stellenbosch

Van der Berg, S., Van Wyk, C., Selkirk, R., & Hofmeyr, H. (2022b). *Learner flows through schools: Using high quality administrative data to understand education system performance*. Resep: Stellenbosch. *https://www.ssrn.com/abstract=4009654*

Van der Berg, S., Van Wyk, C., & Selkirk, R. (2020). Schools *in the time of COVID-19: Possible implications for enrolment, repetition and dropout*. Stellenbosch Economic Working Paper 20/2020. *https://ideas.repec.org/p/sza/wpaper/wpapers327.html*

Van der Berg, S., Van Wyk, C., Selkirk, R., Rich, K., & Deghaye, N. (2019). *The promise of SA-SAMS & DDD data for tracking progression, repetition and drop-out*. Stellenbosch Economic Working Paper 17/2019. *https://www.ssrn.com/abstract=4009654*

Wills, G., & Hofmeyr, H. (2019). Academic resilience in challenging contexts: Evidence from township and rural primary schools in South Africa. *International Journal of Educational Research*, 98(June), 192–205. *https://doi.org/10.1016/j.ijer.2019.08.001*

Appendix tables

TABLE A1 OLS regression: Grade 10 performance

		2019	2020	2021	2022
Gr9 math mark (most recent attempt)		0.644***	0.663***	0.561***	0.624***
Repeated Gr9		–2.077***	0.125	–0.717***	1.00***
Province (base=Gauteng)	EC	–3.065***	–2.580***	0.213*	–1.579***
	LP	3.607***	2.154***	2.267***	–1.216***
Quintile (base=5)	Q1	–11.657***	–4.413***	–8.750***	–8.619***
	Q2	–10.207***	–4.109***	–8.307***	–7.676***
	Q3	–8.167***	–1.682***	–7.349***	–6.366***
	Q4	–7.010***	–2.920***	–6.773***	–6.757***
Female		–0.449***	0.063	0.292***	–0.671***
Overage (base=No)	1yr	6.564*	1.008	–0.583***	–3.838***
	2yr	5.197	0.422	–4.209***	–5.365***
	3+yrs	3.445	–3.333***	–6.135***	–6.326***
Constant		0.766	9.822***	13.757***	9.849***
N		125 828	135 222	152 394	153 607
R^2 (adj.)		0.468	0.411	0.392	0.480
* p<0.05; ** p<0.01; *** p<0.001					

Source: Calculated from DDD data

TABLE A2 Logit regression of Gr10 repetition

		Year of (first) Gr10 attempt		
		2019	2020	2021
Taking Maths in Gr10		–0.710***	–4.05***	–0.267***
Gr10 Maths/MathsLit mark		–0.089***	–0.091***	–0.098***
Gr10 EFAL mark		–0.066***	–0.055***	–0.072***
Province (base=Gauteng)	EC	–0.176***	–0.206***	–0.160***
	LP	0.639***	0.328***	0.341***
Quintile (base=5)	Q1	0.389***	0.419***	0.601***
	Q2	0.522***	0.455***	0.643***
	Q3	0.680***	0.515***	0.734***
	Q4	0.837***	0.610***	0.870***
Female		–0.113***	–0.273***	–0.313***
Overage (base=No)	1yr	–0.680	–1.103***	0.099***
	2yr	–0.731	–1.000***	0.277***
	3+yrs	–0.609	–0.730***	0.180***
Constant		4.966	4.658***	4.351***
N		233 300	234 322	283 85
Pseudo R2		0.300	0.303	0.341
* p<0.05; ** p<0.01; *** p<0.001				

Source: Calculated from DDD data

TABLE A3 Logit regression: On track & passing matric (first attempt)

		2019	2020
Taking maths in gr10 (latest attempt)		0.307***	–0.132***
Gr10 math/math lit mark		0.014***	0.041***
Repeated any year		–2.900***	–0.152***
Province (base=Gauteng)	EC	–0.152***	–0.209***
	LP	–0.139***	–0.445***
Quintile (base=5)	Q1	–0.544***	–0.520***
	Q2	–0.450***	–0.410***
	Q3	–0.372***	–0.508***
	Q4	–0.255***	–0.511***
Female		–0.042***	–0.243***
Overage (base=No)	1yr	0.208***	–0.779***
	2yr	0.515***	–1.212***
	3+yrs	0.348***	–1.639***
Constant		1.635***	1.371***
N		294 555	172 612
Pseudo R^2		0.294	0.110

* p<0.05; ** p<0.01; *** p<0.001

Source: Calculated from DDD data

RESEP
Research on Socio-Economic Policy

Stellenbosch
UNIVERSITY
IYUNIVESITHI
UNIVERSITEIT

forward together
sonke siya phambili
saam vorentoe

RESEP (RESEARCH ON SOCIO-ECONOMIC POLICY)

Department of Economics
Matieland, 7602

(+27) 21 808 2024

carinebrunsdon@sun.ac.za

www.linkedin.com/company/resep-research-on-socio-economic-policy/

@RESEPgroup

https://resep.sun.ac.za/

Michael & Susan Dell
FOUNDATION

We gratefully acknowledge support from the Michael & Susan Dell Foundation