

Pandemic-related losses in contact time across seven provinces according to SA-SAMS data

Martin Gustafsson
22 February 2022

This report uses SA-SAMS data from seven provinces participating in the DDD initiative to address knowledge gaps relating to recent school attendance, in the context of the pandemic. Much of the focus falls on Grade 3, but statistics for grades 6 and 9 are also provided. Data from 2019 to 2021 are analysed, but 2021 Term 3 data are analysed in considerable detail.

There are two key problems with the SA-SAMS attendance data which limit the analysis. Both problems relate to how schools capture absenteeism. Firstly, 12% of schools do not declare anyone as absent in a whole term, which would largely reflect incorrect use of SA-SAMS, as opposed to a reality. It is virtually impossible for no-one to be absent during a term, especially in the context of the pandemic. Secondly, sample-based surveying of schools points to around a third of schools, at the primary level, employing rotational attendance to promote social distancing, yet SA-SAMS data point to just 12% of primary schools having this, and perhaps as few as 5% entering details in this regard correctly on the system.

In the case of the relatively few schools where data on rotations exist, one finds that around 20% of primary schools with rotations follow weekly rotations, with learners attending every second week. For the remaining 80% of schools with rotations, some form of rotation within the week occurs. At the secondary level, the figure for weekly rotations is a bit higher, at 30%.

The key question addressed in this report is how many schools which do *not* display rotations in their SA-SAMS data do in fact follow this practice. This is a vital question for understanding the extent to which contact time has been lost. The SA-SAMS data can to some extent provide answers insofar as patterns of *regular* absenteeism do, to some extent, reflect rotation patterns. A learner who is only permitted to attend on 50% of school days, due to rotations, can only be marked as being regularly absent on such a day. This means that if there is enough regular absenteeism occurring in a school, and if this regular absenteeism follows a pattern that is compatible with some rotation pattern, then the likelihood that the school practices a rotation rises.

The finding is that patterns of regular attendance *do* display evidence of rotations, and that an indicator of the likelihood that rotations are being followed can be constructed for every school with relatively high levels of regular absenteeism. This indicator suggests that at the primary level, between 20% and 50% of schools did rotations. The percentage is probably not lower than this range, and probably not higher than it. This would be in line with other evidence that around a third of schools practiced rotations, at least in 2021 Term 3.

The key national ‘attendance percentage’ statistics are as follows.

- The ‘at face value’ attendance percentage for 2021 Term 3 Grade 3, drawing from SA-SAMS, is 90%. This is what the unadjusted SA-SAMS data point to. On average, Grade 3 learners lost 10% of contact time. The 90% statistic represents a six percentage point decline relative to the pre-pandemic situation. Declines of a similar magnitude are seen for grades 6 and 9. Such a decline is worrying. Yet this ‘at face

value' decline *under-states* the problem due to the two data entry problems mentioned above.

- If one makes adjustments to take into account the fact that some schools did not even enter *regular* (non-rotational) absenteeism into SA-SAMS, and the fact that many schools did not enter data on rotations, even though they had rotations, the 90% attendance percentage for Grade 3 drops to 78% across the seven provinces. This is if evidence from outside SA-SAMS that around a third of schools practiced rotations in Term 3 of 2021 is used to find a point within the abovementioned range of 20% and 50% of schools. This range of 20% to 50% of schools implementing rotations translates to a range in the attendance percentage of 72% to 86%. It turns out that the non-entry of regular absenteeism biases the statistics only marginally. The greatest distortion is caused by the non-entry of information on rotations.
- The estimate that 22% of contact time was lost during Term 3 of 2021 in Grade 3 (100% minus 78%) can be deceptive as this average hides the fact that losses have been highly unequal. It hides the fact that just over a third of learners in Grade 3, but also similar numbers of learners in other grades, lost as much as 50% of contact time. This was the situation for learners in schools with rotations. The SA-SAMS data, and other data, point to rotational arrangements nearly always being about dividing the school into two groups, not more than two groups. The effect on the statistics of taking into account the extremely few schools with more than two groups is negligible.
- In many provinces, the gap between the 'at face value' attendance percentage and the percentage after adjustments is very large. The data in SA-SAMS are particularly bad at reflecting the problem of losses in contact time in KwaZulu-Natal, Limpopo and North West.

The report ends by comparing enrolments at the start of 2021 to attendance in Term 3 of 2021. In an earlier report, enrolments at the start of 2020, just before the pandemic, had been compared to enrolments at the start of 2021. The finding was that enrolments at the start of 2021 were around 50,000 lower than one might expect. This could plausibly be attributed to the pandemic. A key question, however, was whether 2021 enrolments were inflated by schools to protect school resources, implying a decline in enrolments that would be larger than 50,000. The attendance data analysed for this report provided an opportunity to test whether this did occur. Attendance data, unlike enrolment data, are not used for resourcing, so there is a much smaller incentive to manipulate it for fraudulent purposes. The analysis presented below, which compared exactly the same schools across the enrolment and attendance data for 2021, points to no major attrition between the enrolment and attendance data. In grades 3 and 6 learner numbers match almost perfectly, while for Grade 9 a decline of 1.1% is found. The latter difference could be the result of dropping out in Grade 9 during the 2021 school year, but what could not be ascertained with the available data is whether this is unlike the pattern before the pandemic. At the Grade 9 level, learners are allowed to leave school. What the analysis shows, when combined with the findings from the earlier enrolments report, is that it seems extremely unlikely that dropping out increased by half a million or more as a result of the pandemic. This was widely reported in the media in the middle of 2021, following a telephonic household survey.

Contents

1	Introduction	3
2	Other evidence on COVID-related losses in contact time	4
3	The SA-SAMS data on attendance	5
4	Attendance aggregates over time in seven provinces ‘at face value’	6
5	First problem: schools where no learner is ever absent	10
6	Second problem: rotations not marked as missed school days	11
6.1	Rotation patterns where school data are relatively good	11
6.2	Detecting rotation patterns in schools without good data.....	14
7	Arriving at plausible adjusted statistics on the loss of contact time	19
8	Was attendance lower than enrolments in 2021?	20
8.1	The possibility of ‘ghost learners’ with 100% attendance.....	20
8.2	A comparison of the 2021 enrolments and attendance data	21

1 Introduction

This report draws from a large dataset extracted from the Data Driven Districts (DDD) dashboard, also known as the ‘DBE Dashboard’¹. These data ultimately draw from the SA-SAMS² tool used in schools.

The report assists in filling crucial information gaps relating to losses in contact time among learners due to the pandemic. Learning losses are beginning to become clearer, drawing from test data before and after the pandemic³. Learning losses are to a large extent driven by declines in contact time, whose magnitude and nature is still poorly understood. While the attendance data made available for this report has limitations and is not able to answer all questions, it remains a valuable source of information. Moreover, while the analysis presented below is detailed, it is possible that there are ways of gauging rotational attendance in schools with imperfect data beyond what is explored here.

The aim of the report is both to produce statistics on losses in contact time, and to suggest how existing data collection systems could be strengthened.

In this report, **losses in contact time** means losses in daily school attendance brought about by the pandemic. This can be of two types. Firstly, this can be due to lower than usual levels of attendance *on days when the learner is permitted to come to school*. Secondly, it can be due to *days when only some learners, but not others are permitted to come to school*, a restriction imposed in order to enhance social distancing. As will be seen in the report, this distinction is vital if the data are to be analysed correctly. What is not within the scope of the report is the impact of the pandemic on time use once the learner is in the school. In other words, attending school for a day before the pandemic is considered the same as attending school for a day during the pandemic. This is obviously simplistic, but the data would not permit an analysis of this aspect of time use.

The data extract made available for the analysis allowed for grades 3, 6 and 9 to be analysed. In several parts of the report the focus falls specifically on Grade 3, in part due to the especially strong need for daily attendance at this level, given that foundational skills are

¹ <https://dbedashboard.co.za>.

² South African School Administration and Management System.

³ See for instance Shepherd, Mohohlwane, Taylor and Kotzé (2021) and Ardington, Wills and Kotze (2021).

being developed. However, as will be seen, Grade 3 patterns are fairly representative of patterns seen in the other two grades.

2 Other evidence on COVID-related losses in contact time

This report represents the first time that data one could consider more or less system-wide administrative data are being used to gauge *attendance*. *Enrolments* are far more commonly analysed using administrative data. For instance, a report on enrolments at the start of 2020 compared to the start of 2021 was released by the DBE in October 2021 to gauge the impact of the pandemic on 2021 enrolments⁴. Enrolments are of course registration with a school as a learner, while attendance is actual daily participation.

The monitoring of attendance has so far occurred through the use of sample-based household data and sample-based collections from schools.

Household data from the NIDS-CRAM initiative resulted in a paper by Shepherd and Mohohlwane (2021), which concluded that participation in schooling in April 2021 was around half a million lower than it would have been without the pandemic. This half a million translates to around **a four percentage point drop in participation**. This figure refers to the situation only on days when the learner is permitted to attend school. In other words, the figure **under-states the problem insofar as it does not take into account widespread rotational attendance arrangements introduced during the pandemic**. Furthermore, it should be noted that the focus of the NIDS-CRAM report is on any participation over a longer period of time, and not the regularity of attendance.

The NIDS-CRAM analysis has been revised by the authors of the original paper and will be released both as a journal article and a media article. The author of the current paper participated in the re-evaluation of the original analysis⁵. The re-analysis confirmed that the participation problem is a serious one, with the drop in participation probably being above half a million, possibly in the region of 900,000. Very importantly, the re-analysis found that the problem was far more serious at the primary level than at the secondary level.

In November 2020, a group of reading intervention primary schools in Mpumalanga were asked about their rotations. 16% of schools indicated they allowed learners to come to school every day. Among the remainder, around three-quarters ran within-week rotations, meaning learners attended some days each week. A quarter used across-week rotations, meaning learners would attend every second week⁶. **If around 80% of learners are allowed to attend school only half of the time, this implies ‘absenteeism’, in inverted commas because the learner is not permitted to attend school on some days, of 40%**⁷. 40% of potential contact days are lost, *and this ignores absence on days when learners are permitted to attend*.

Data collected from 120 Limpopo no-fee primary schools during 2021 paint a picture similar to the Mpumalanga 2020 one, at least for terms 1 and 2 of 2021: 10% of schools in the first half of 2021 had attendance for everyone every day, and for two-thirds of the remaining schools it was within-week rotations. However, the situation had improved considerably by Term 3 of 2021, according to the same dataset: at that point 48% of schools had attendance for everyone every day⁸. The percentage of schools where learners could attend school *every third day*, as opposed to every second day, was 8% in terms 1 and 2 of 2021, and 4% in Term 3.

⁴ Department of Basic Education, 2021.

⁵ Internal DBE paper by Gustafsson titled ‘By how much has school participation declined as a result of the pandemic?’, date 30 November 2021.

⁶ Shepherd *et al*, 2021: 27.

⁷ 80% of learners attend half the time, 20% of learner attend every day.

⁸ Excel spreadsheet from Cally Ardington.

In its interactions with provincial counterparts, the DBE has requested information on rotational attendance. Information from these interactions, compiled in a slide presentation⁹, includes attendance per day and per grade per province, for the week starting 16 July 2021. All provinces except Free State are included. To illustrate, Grade 3 attendance rates across the five days range from an average of between 60% and 70% for Gauteng and KwaZulu-Natal, to over 90% for Limpopo and Mpumalanga. The range suggests that in some provinces absence during a rotation off-day was counted as an absence, while in others it was not.

A December 2021 presentation of the National Education Collaboration Trust (NECT) provides statistics collected from a national sample of 265 primary schools in terms 3 and 4 of 2021¹⁰. Statistics point to **35% of schools allowing learners to attend every day, with most of the remainder practicing within-week rotations**. Moreover, responses from school principals in the sampled schools suggest 10% of schools planned to end rotations in early 2022, which would reduce the 35% to 25%.

Clearly, statistics vary considerably, and often focus on different aspects of the problem. The assumptions in the analysis that follows will draw from the NECT survey, which has the advantage of being both national and recent.

3 The SA-SAMS data on attendance

The statistics in the DBE Dashboard are viewable only by registered users. Statistics in the dashboard, down to the level of schools and anonymised learners, were viewed to gain a sense of how attendance is covered. This led to a data request, which is reproduced in Annexure 1 below.

The number of schools for which attendance data were available in each term is shown in Table 1. Except for terms 2 to 4 of 2020, which coincided with the start of the pandemic, the numbers are consistent and high, and never below 20,000.

Table 1: Schools with DDD attendance data 2019-2021

Year	Term	Schools with attendance data (7 provinces)
2019	2	21,271
2019	3	20,965
2019	4	20,573
2020	1	20,003
2020	2	3,972
2020	3	17,472
2020	4	19,762
2021	1	20,759
2021	2	20,396
2021	3	20,134

Provincial details are provided in Table 2. Just over a thousand schools are independent. Only in Gauteng are fewer than 99% of public schools presents at some point in time, within the period 2019 Term 2 to 2021 Term 3. The final column indicates the number of schools present in seven periods, these being all periods in Table 1 other than the three problematic 2020 terms. The 16,313 schools are used for much of the comparison over time that appears below.

⁹ Titled ‘DG’s progress report’ and dated 10 September 2021.

¹⁰ Title ‘COVID Response Initiative data collection’, dated 15 December 2021.

Table 2: Details on numbers of DDD schools

	Schools with attendance data in any year (A)	Number from A which are public (B)	Official 2019 number of public schools (C)	% B over C	From B with all 7 term values outside 2020 T2 to 2020 T4
EC	5,413	5,187	5,205	100	4,893
GP	2,846	2,225	2,071	93	1,654
KN	6,076	5,895	5,821	99	2,966
LP	3,945	3,786	3,773	100	3,596
MP	1,790	1,678	1,679	100	1,418
NC	578	550	546	99	503
NW	1,537	1,463	1,451	99	1,283
Total	22,185	20,784	20,546	99	16,313

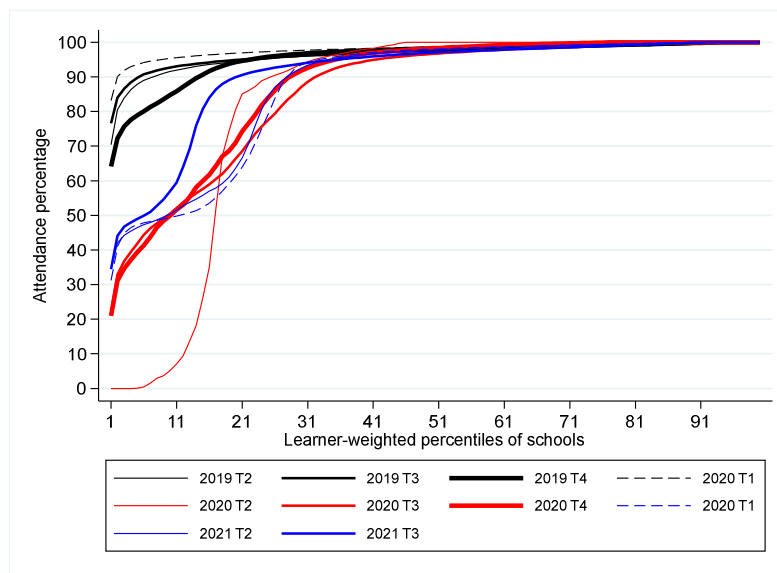
Learner coverage within schools seems to be very high, and complete. To illustrate, 794,776 Grade 3 learners are covered in the 2021 Term 3 data.

4 Attendance aggregates over time in seven provinces ‘at face value’

This section presents what the data say ‘at face value’. For two key reasons, this under-states actual absenteeism. Firstly, some schools appear never to report anyone as absent in a term, which is virtually impossible. Secondly, it is clear that some schools classify learners as being present, or non-absent, when they are forced to remain at home due to a rotation. The extent of these problems is analysed in sections 5 and 6. Yet even if ‘at face value’ statistics are an under-statement, they are a good point of departure, and provide an indication of the *best case scenario*.

Figure 1 illustrates the distribution of attendance across 12,090 public schools, these being all the 16,313 with enough data over time which also have Grade 3. The ‘at face value’ distributions suggest that for around two-thirds of Grade 3 learners attendance remained over 90% even during the pandemic. However, for a third of these learners attendance dropped drastically. **The aggregate attendance statistic represented by the 2019 Term 3 curve is 97%, while that for 2021 Term 3 is 91%. Even a decline of six percentage points is very concerning, but the actual situation is worse than this.**

Figure 1: Grade 3 distribution of absenteeism over time



The following two graphs provide a similar analysis for grades 6 and 9. The picture is not that different to the Grade 3 picture. **The aggregate decline in attendance across the two years for Grade 6 is from 97% to 90% and that for Grade 9 from 95% to 88%** (for each year Term 3 is considered).

Figure 2: Grade 6 distribution of absenteeism over time

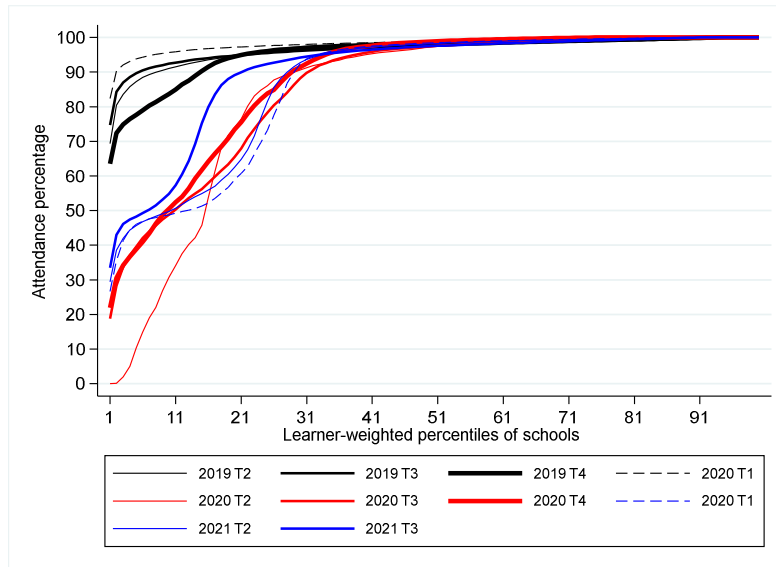


Figure 3: Grade 9 distribution of absenteeism over time

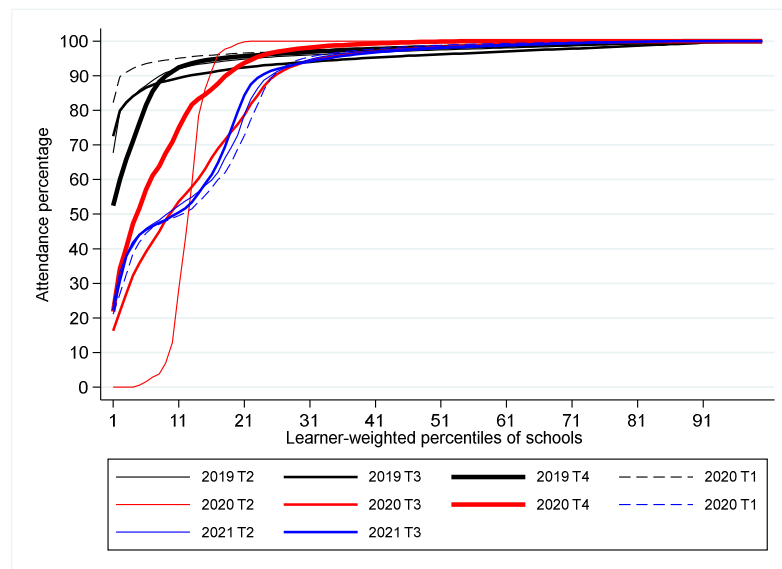
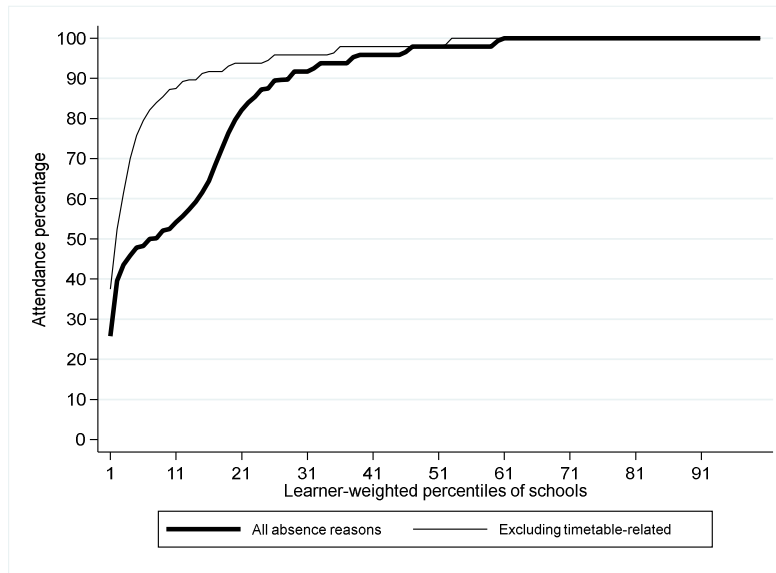


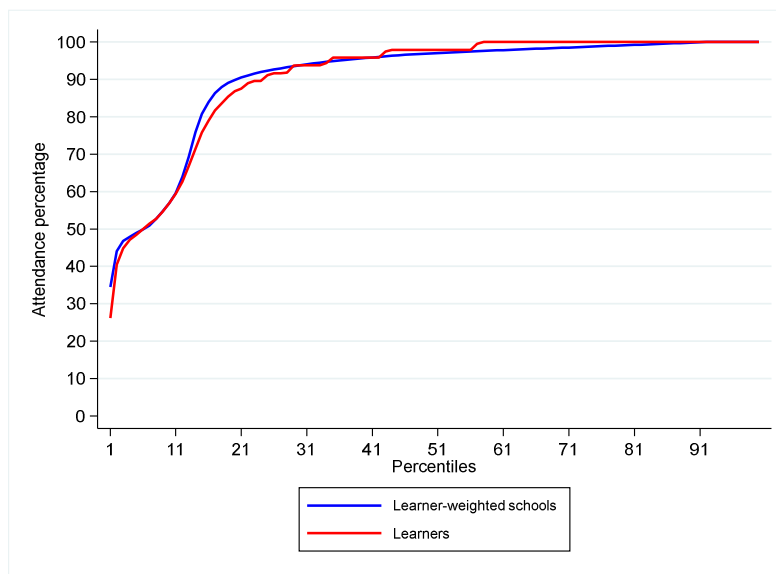
Figure 4 below points to a critical feature of the data. Some schools, but clearly not enough schools, are using a new feature in SA-SAMS. That new feature is the following reason for a learner absence: 'Not attending school as per timetable'. This is designed for a situation where rotational attendance is implemented and there are days when the learner is absent essentially because the learner is not permitted to attend school, because it is not the turn of the learner. If this category is ignored, in other words if one focusses only on attendance *when a learner is permitted to attend school*, the statistics improve vastly. Then around 80% of learners, in the case of Grade 3, get to attend at least 90% of the time.

Figure 4: 2021 Term 3 Grade 3 distributions with and without rotation



The above graphs use learner-weighted school values, which could hide extreme differences within schools, and hence for instance a high number of learners with attendance at around 10%. Figure 5, which focusses on Grade 3, suggests this is not the case. Whether learner-weighted schools or learners is the unit of analysis, the distribution looks very similar.

Figure 5: 2021 Term 3 Grade 3 distributions for schools and learners



There is hardly any difference across gender for the 2021 Term 3 statistics. For instance, for both grades 3 and 9 female attendance exceeds male attendance by just 0.2 percentage points.

The following three graphs break the trend over time up by province, quintile and whether a school is independent. Up to Figure 7, all the analysis focusses on just public schools. The 'at face value' level of attendance in Gauteng and Northern Cape remained worryingly low, at around 80%, even in 2021 Term 3. And this is the most optimistic outlook possible. The

pattern across quintiles is systematic: the poorer the school, the less contact lost, according to this view of the data. Independent schools appear to have incurred losses in contact time which are lower, relative to public schools, by around three to four percentage points.

Figure 6: Provincial Grade 3 attendance trends over time

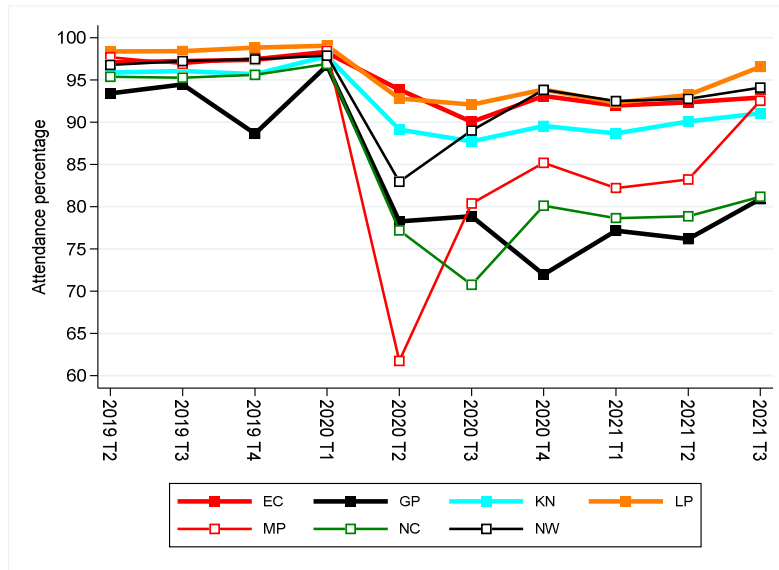


Figure 7: Quintile-specific Grade 3 attendance trends over time

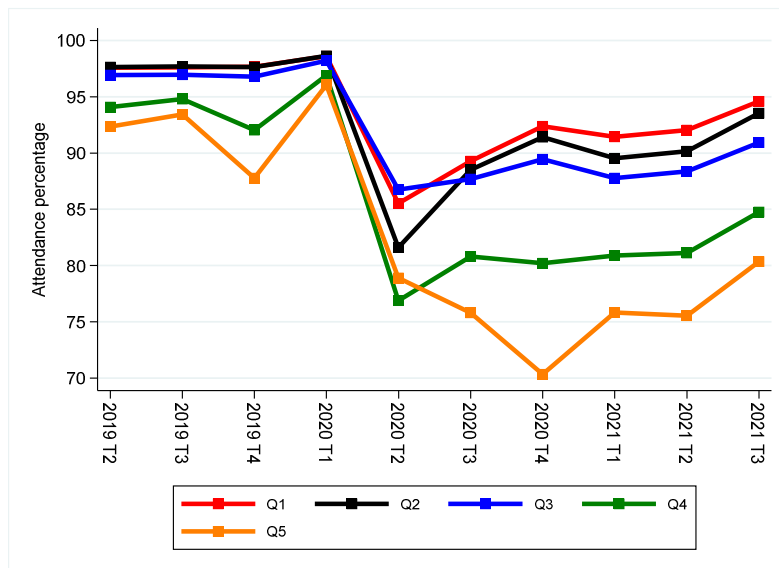


Figure 8: Grade 3 attendance trends over time by sector

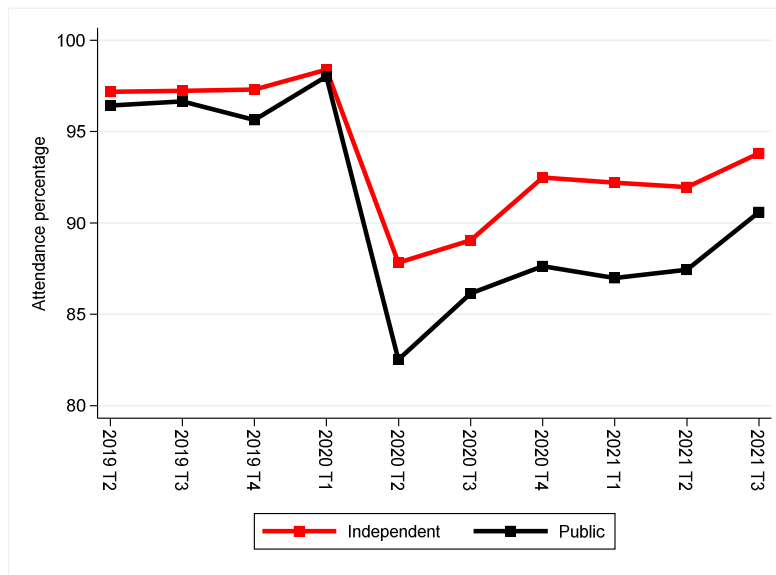
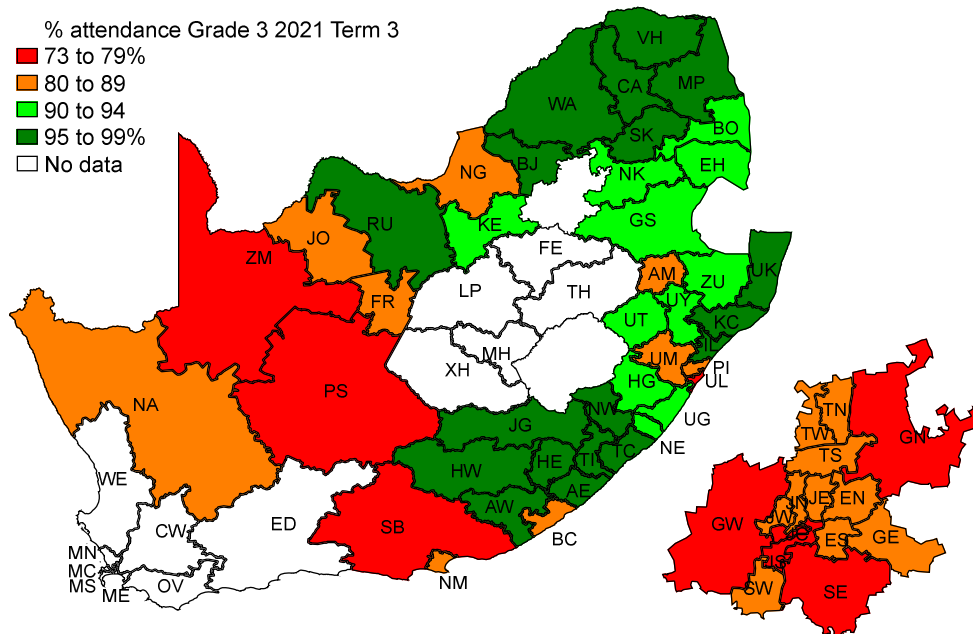


Figure 9 below uses Grade 3 data from 13,728 public schools appearing in the 2021 Term 3 data which could be linked to district, via the schools master list, regardless of whether the schools appear in earlier time periods in the SA-SAMS data. The ‘at face value’ statistics point to considerable within-province inequality in the case of KwaZulu-Natal, Eastern Cape and Northern Cape. In these three provinces, districts display the full range of attendance levels.

Figure 9: 2021 Term 3 Grade 3 attendance by district



5 First problem: schools where no learner is ever absent

As shown in Table 3, in 12% of schools not a single learner was found to be absent, in any of the three grades, during 2021 Term 3. This was most prominent in Eastern Cape and KwaZulu-Natal. It can be considered extremely unlikely that no learner would be absent,

especially during a pandemic with health and economic crises occurring in households, and schools under the obligation to keep learners displaying flu-like symptoms out of the school for a period. It can be assumed that virtually all the 2,261 schools with no absence data simply did not enter absent learners into SA-SAMS. Clearly, if more realistic aggregate attendance statistics are to be derived, an adjustment would have to be made for this anomaly. At the same time, excluding schools with no absenteeism data does not make a large difference to the statistics presented in section 4. For example, the overall Grade 3 attendance level drops by just 0.7 percentage points.

Table 3: No absent learners in 2021 Term 3

	Public schools with attendance data in any of three grades (A)	Number from A with no absent in any grade 2021 Term 3 (B)	% A over B
EC	5,030	904	18
GP	2,061	33	2
KN	4,917	742	15
LP	3,675	335	9
MP	1,606	149	9
NC	535	6	1
NW	1,395	92	7
Total	19,219	2,261	12

6 Second problem: rotations not marked as missed school days

6.1 Rotation patterns where school data are relatively good

Of the 13,976 public schools with Grade 3 attendance data for 2021 Term 3, 1,714 schools were using the new ‘Not attending school as per timetable’ attendance reason. Relevant school counts appear in Table 4. Clearly, a high percentage of Gauteng’s schools indicated via SA-SAMS that they had rotational arrangements. This is likely to explain this province’s exceptionally low attendance in earlier Figure 5. The 1,714 schools were fairly equally spread across the five school quintiles.

Table 4: Use of the rotation classification for Grade 3

	Public schools with Grade 3 data (A)	Schools from A using the rotation category (B)	% A over B
EC	4,184	201	5
GP	1,415	490	35
KN	3,549	370	10
LP	2,341	210	9
MP	1,098	237	22
NC	395	118	30
NW	994	88	9
Total	13,976	1,714	12

An analysis was undertaken of the 1,714 schools using the new classification of absence. The aim of this was to examine what the typical rotation patterns are, in part because this might assist in detecting unreported rotations in the over 10,000 schools with Grade 3 not reporting this.

The approach was as follows. A rotation pattern was considered any attendance pattern an individual learner followed, where only days of ‘Not attending school as per timetable’ were counted as days absent. There were 48 school days in 2021 Term 3, spread over ten weeks, with two public holidays automatically considered non-school days. Any pattern involving the learner’s loss of 12 or more school days was considered. Moreover, for a pattern to be included in the analysis, at least 20 learners in the school’s Grade 3 level each had to lose at

least 12 days. Thereafter, any pattern of on and off days shared by at least 15 learners was considered a pattern of interest. A school might have one, two or more such patterns. For instance, if each learner attended every second week, there would be two patterns, and every learner would lose half of the 48 days. Of the 1,714 schools, 637 schools displayed patterns which passed the criteria mentioned above. In the remainder of schools, often learners lost too little time, or there were many patterns with few learners following each pattern. Examination of the data suggests that there were schools which entered the ‘Not attending school as per timetable’ on a few occasions by mistake. But it also seems possible that many schools which were applying rotations, and were attempting to enter this on SA-SAMS, did not do this correctly.

Of the 637 schools with patterns of interest, 150 displayed just one pattern of interest. For instance, some learners would attend every second week, but then other learners would attend almost all the time, or be spread across a multitude of less usual patterns. This left 487 schools with more than one pattern of interest each. Of the 487, 327 displayed two patterns and 160 had more than two patterns. The latter would be possible if for instance a school divided learners into three groups and had each group attend on different days.

Among the 487 schools, the most common two-pattern arrangement involved dividing learners into two groups and having each group attend every second day. This would mean one group’s school week would begin on Monday one week, and on Tuesday the next week. This was seen in 25 schools. A further 22 schools displayed an almost identical pattern, the difference being that groups switched the day after the 9 August public holiday. A further 13 schools displayed two groups, each with attendance every second *week*, for the whole week. And so on. Even small differences relating to one day would result in a different pattern being identified. Hence the result of the analysis was not the identification of patterns shared by large numbers of schools. In total, there were 935 different patterns found across the 637 schools. Many patterns were unique to specific schools, or shared across just one or two schools.

To facilitate the analysis, the 935 patterns were characterised by whether they contained any of six weekly patterns in the eight weeks without public holidays. The six weekly patterns are represented by column headings in Table 5. To illustrate, ‘YYYYY’ means attendance from Monday to Friday, and ‘YNYNY’ means attendance on just Monday, Wednesday and Friday. ‘YY-NN’ means attendance on Monday and Tuesday, and a timetable-driven absence on Thursday and Friday, with what happens on Wednesday being ignored. The final column reflects the number of schools experiencing the combinations of the weekly patterns. The full table would reflect 32 rows. The sum of the final column across the 32 rows was 762, though the analysis covered just 637 schools. This is because one school could be counted across more than one row, as a school could have more than one pattern, with patterns not necessarily belonging to the same group (a row in Figure 5 represents a *group* of patterns).

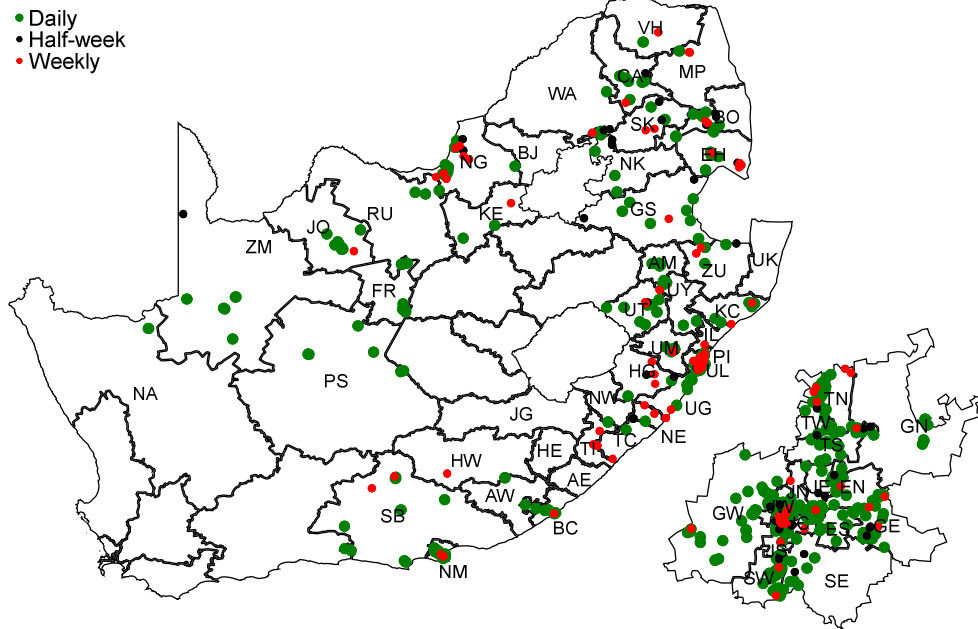
Table 5: Rotation patterns categorised

YYYYY	NNNNN	YNYNY	NYNYN	YY-NN	NN-YY	Schools
0	0	1	1	0	0	220
1	0	1	1	0	0	210
1	1	0	0	0	0	91
1	0	0	0	1	1	32
1	0	0	0	1	0	18
0	0	0	0	0	1	17
1	0	0	0	0	1	15

Further categorisation was undertaken. Any school with a value 1 in *both* of the first two columns of Table 1 was placed in a first category, any school with a value 1 in *both* of the middle two columns was placed in a second category, and any school with a value of 1 in *both* of the last two columns was placed in a third category. These three categories could be

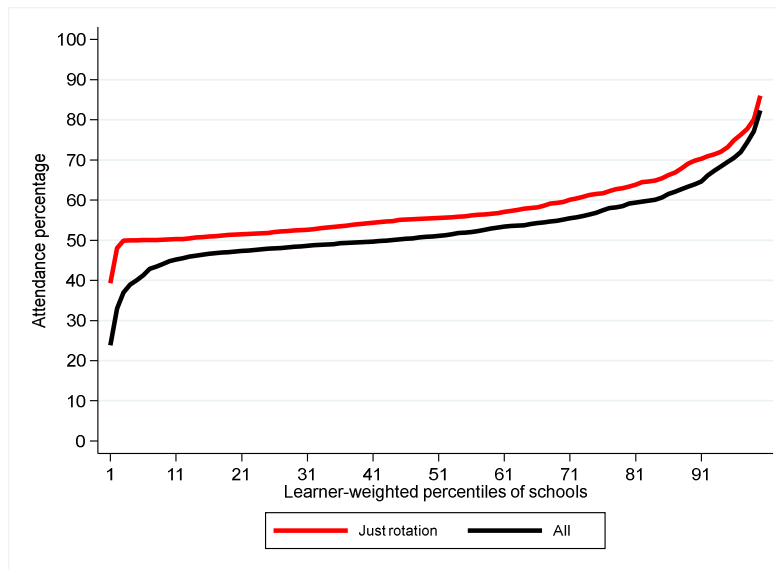
labelled ‘weekly’, ‘daily’, and ‘half-week’. Of the 637 schools, 559 were located in just one of the three categories. An attempt was made to map all these 559 schools – geo-coordinates were found for 553 of them, and these appear in the map of Figure 10 below.

Figure 10: 2021 Term 3 Grade 3 rotations for 553 schools with clear data



The distribution of attendance and its complement, absenteeism, in the 559 schools is shown in Figure 11 below. **99% of learners experience an attendance level of 50% or more as a result of just rotations, suggesting that dividing schools into more than two attendance groups is very rare.** The fact that attendance slightly above 50% is not uncommon would be in part be driven by the fact that rotational arrangements were applied in just certain weeks, and not all weeks. If absenteeism not related to rotations is also taken into account, attendance drops by around three percentage points in the 559 schools.

Figure 11: Grade 3 distribution of attendance in 559 schools with clear rotation data



Disaggregating the 559 schools by quintile and type of rotation produces the following table – the total number of schools is 551 as these are the schools for which a quintile could be found in the master list of schools. **Three-quarters of schools with Grade 3 and with rotation data practiced daily rotations, as opposed to half-week and weekly rotations.**

Table 6: Grade 3 schools by quintile and type of pattern

Quintile	Daily	Half-week	Weekly	Total
1	47	15	19	81
2	67	11	15	93
3	93	9	29	131
4	75	4	17	96
5	126	3	21	150
Total	408	42	101	551

The corresponding tables for grades 6 and 9 appear below. For these tables, the entire process described for Grade 3 above was repeated. On the whole, the patterns across grades are similar, though **in Grade 9 weekly rotations feature somewhat more prominently.**

Table 7: Grade 6 schools by quintile and type of pattern

Quintile	Daily	Half-week	Weekly	Total
1	64	9	13	86
2	88	8	21	117
3	126	17	27	170
4	94	6	18	118
5	136	4	20	160
Total	508	44	99	651

Table 8: Grade 9 schools by quintile and type of pattern

Quintile	Daily	Half-week	Weekly	Total
1	46	14	32	92
2	55	17	39	111
3	61	11	55	127
4	46	10	14	70
5	70	1	12	83
Total	278	53	152	483

6.2 Detecting rotation patterns in schools without good data

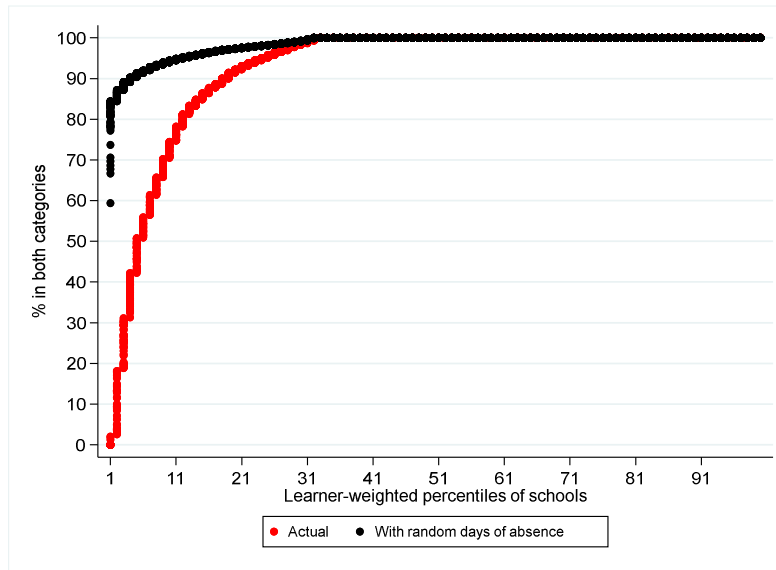
In theory, it seems possible to gauge whether schools are practicing rotations or not by examining just regular absenteeism, in other words not absenteeism ‘as per timetable’. In a simple setup where only one rotation system was followed, and where there were many learners experiencing regular absenteeism on more than one day in the term, one would find that learners who were absent for several days would divide neatly into two groups. One group would experience absence only on the 50% of days when they were expected to attend, and a second group would only be absent during the other 50% of days. No learner who is absent more than once would have such days spread across the two sets of days.

However, the setup is not neat, for two key reasons. Firstly, many different patterns of rotation are found. Secondly, in many schools there are too few learners absent on more than one day, and preferably on many days, to make the analysis feasible. The aim of what follows is to attempt to detect patterns despite these factors.

Both curves in Figure 12 draw from the 4,820 public schools with Grade 3 which remained after two filters were applied. First, any school where no learner experienced five or more

days of absenteeism was dropped. The second filter, and it was applied after the first filter had already been applied, was to drop any school with fewer than 30 days absent, across all learners, in Grade 3.

Figure 12: Possible effects of rotations on regular absenteeism



For Figure 12, just one rotation pattern was tested. This was the rotation pattern that emerged as the most common, and was found in 25 schools – see the discussion in section 6.2. This was a very specific daily rotation system, with learners attending every second day. Using this pattern, days of absenteeism for each of the 4,820 schools could be divided into three groups. Firstly, there would be days where the same learner’s days all fell into days comprising the pattern. Secondly, there would be days where same learner’s days all fell outside the set of days comprising the pattern. Thirdly, there would be days absent where the same learner’s days spanned both sets of days.

The red curve in the graph captures the percentage of days of absenteeism per school, after the filters had been applied, which fell into the third group, in other words days which did not follow the rotational pattern being analysed. Only 36 schools had no learners in this third group. These *could* be schools which followed the pattern in question. But it is also possible that they did not, but days absent followed the pattern by accident. It is also possible that schools with a value greater than zero in the third category *did* deliberately follow the pattern, but that some learner fell into the ‘both’ category due to errors in data entry.

The black curve in the graph was produced by using the 4,820 schools, and taking the same *number* of days of absenteeism used for the red curve. The difference was that for each learner, the days absent were rearranged randomly over the 48 school days of the term. The vertical gap between the two curves suggests that for up to around 30% of schools there is a *possibility* that a low value in the ‘both’ group in 2021 Term 3 is due to the fact that schools followed a rotation like the pattern in question, or similar to it.

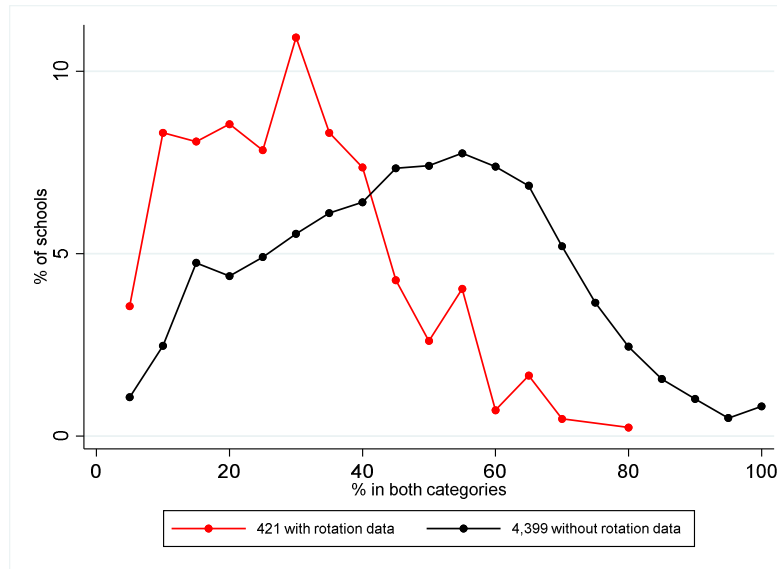
The remainder of this section will explain what one finds if one repeats the calculations behind the red curve for all the patterns seen in the 637 schools where we *do* have relatively good data on rotations. Thereafter, how the results of these calculations, which draw from 4,820 schools, were compared to patterns seen in 421 of the 637 schools, is discussed. The 421 schools constitute the subset of the 637 schools where there is enough data on *regular*

non-timetable absenteeism for the comparison to be possible. In other words, the 421 schools are a subset of the 4,820 schools remaining after the filters were applied.

The number of patterns seen in the 637 schools is 935, as discussed in section 6.1. For each of the 4,820 schools, the alignment between regular days of absenteeism and the pattern were assessed, the result being a ‘% in both categories’ for each calculation. There were around 4.5 million calculations – 4,820 times 935. Figure 13 illustrates the *minimum* of the 935 ‘% in both categories’ per school. In theory, this would represent the statistic for the rotation pattern that most fits the pattern of regular absenteeism. What is not illustrated in the graph is that exactly 537 schools among the 4,339 display exactly zero: all days absent fit perfectly with at least one rotation pattern. 537 over 4,339 is 12%. The corresponding values for the 421 schools with rotations is 91 schools and 22%.

It is clear that the 421 schools which we know have rotations display relatively low ‘% in both categories’ statistics, as one might expect. The question is what cut-off on the horizontal axis could be used for the schools without rotation data to estimate, very roughly, which schools have rotations. It could be argued the cut-off should be 15%. There is a noticeable minor peak in the black curve, suggesting rotations are playing a role at this point. Or it could be argued that the cut-off should be around 40% – schools with rotations have statistics clustered below around 40%.

Figure 13: Identification of possible rotations across 4,820 schools



Note: For the horizontal axis, values were rounded to the nearest 5%. Data with 0% on the horizontal, meaning greater than 0 and less 2.5%, is not illustrated.

If the black curve in Figure 13 is limited to just quintiles 1 to 3 schools, or just quintiles 4 to 5 schools, the general form of the black curve does not change. This suggests that rotations were not concentrated in schools serving poorer communities, or schools serving better off communities.

If the NECT sample-based estimate of 35% of all primary schools practicing rotations is used – see section 2 – then the ‘% in both categories’ cut-off would be 30%. This considers the data from all the 4,820 schools at once. The aforementioned cut-off of 15% would correspond to 50% of learner-weighted schools, while a cut of 40% would correspond to 21% of learner-weighted schools. Four randomly selected schools emerging when a cut-off of 20% is used

are reflected in Table 9. These are schools which did *not* use the timetable reason for absence on SA-SAMS. The selection of the four schools occurred by selecting the first schools appearing in an initial search, the only criterion being that they had to be from four different provinces. No selection was rejected because it did not seem to reflect a clear rotation pattern. The aim here is thus for Table 9 to reflect what is actually in the data. It is clear that the methodology used is not 'watertight'.

The second school in Table 9, in Eastern Cape, appears to be a clear case of rotations being employed, though this is not specified in SA-SAMS using the timetable reason. In this school, the pattern that fits the regular absenteeism data is a clear case of a daily rotation. The two colours illustrate the pattern followed. Moreover, regular absenteeism seems clearly divisible into two distinct groups of learners in this school, corresponding to the two rotation groups, referred to as A and B here. In the case of the other three schools, the conclusions are less obvious. Only the second and fourth schools, from Eastern Cape and Mpumalanga, have days absent from high-absence learners spread across both groups A and B. In the other two schools, days absent could be found that followed the pattern, or A, but not the inverse of the pattern, or B. Among the 4,399 schools with no timetable-linked days of absenteeism, 83% have days absent following just one pattern. The figure is 53% for the 421 schools with timetable-linked absences. Thus even for schools which clearly did have rotations, the correspondence between regular absenteeism and rotation is not straightforward.

Table 9: Daily Grade 3 details for four randomly selected schools

School EMIS number:		922221531 (LP)			200200677 (EC)			500323084 (KN)			800013300 (MP)		
		A	B	Both	A	B	Both	A	B	Both	A	B	Both
26-Jul-21	Mon				4		2	6			1		
27-Jul-21	Tue					2	2	4			1		
28-Jul-21	Wed				3			3			1		
29-Jul-21	Thu					2	2	2			1		
30-Jul-21	Fri				6						1		
02-Aug-21	Mon					5		1			3		
03-Aug-21	Tue				3			4			4		
04-Aug-21	Wed					2					4		
05-Aug-21	Thu				3						4		
06-Aug-21	Fri					4					3		
10-Aug-21	Tue				11		2				3		
11-Aug-21	Wed					4	1	2		1	3		
12-Aug-21	Thu				7		1				2		
13-Aug-21	Fri					4	3				3		
16-Aug-21	Mon			2								1	
17-Aug-21	Tue			2				2				1	
18-Aug-21	Wed			2	15		1					1	
19-Aug-21	Thu	5				5	2					1	
20-Aug-21	Fri	5			12		1					1	
23-Aug-21	Mon	8				8	3	1			3		
24-Aug-21	Tue	8			11		1				2		
25-Aug-21	Wed	8				7	2						
26-Aug-21	Thu			2	10		1			1	1		
27-Aug-21	Fri			2		8	1	2					
30-Aug-21	Mon	5			6								
31-Aug-21	Tue	5				1	1						
01-Sep-21	Wed	5			6		1						
02-Sep-21	Thu					1	1						
03-Sep-21	Fri				5								
06-Sep-21	Mon	1				4		1		1			
07-Sep-21	Tue	1			7		2	7		1	1		
08-Sep-21	Wed	5				2	2	3					
09-Sep-21	Thu	5			3								
10-Sep-21	Fri	5				3	1						
13-Sep-21	Mon	2			7								
14-Sep-21	Tue	2				4							
15-Sep-21	Wed	3			5								
16-Sep-21	Thu	3				2	1						
17-Sep-21	Fri	3			8								
20-Sep-21	Mon	3		1				4					
21-Sep-21	Tue	3			2		1						
22-Sep-21	Wed	3		1									
23-Sep-21	Thu	2			2								
27-Sep-21	Mon	14		1				2					1
28-Sep-21	Tue	14		1				1		1			1
29-Sep-21	Wed	17		2									1
30-Sep-21	Thu	13		3				3					1
01-Oct-21	Fri	12		3									1

The use of regular attendance to detect rotations is clearly not sufficient to identify reliably *individual* schools that use rotations. The methodology is, however, arguably useful for identifying which provinces have implemented rotations more extensively than others. It is likely that rotations would drive the ‘% in both categories’ statistics in similar ways in different provinces.

Table 10 indicates the percentage of Grade 3 learners in schools employing rotations in 2021 Term 3, according to the timetable-related data and according to the methodology presented above. The latter assumes the NECT conclusion that 35% of *schools* (not necessarily learners) experienced rotations in this term. Moreover, it is assumed that each province’s situation is

fairly accurately illustrated by the province's share of the 4,820 schools nationally which had relatively high levels of regular absenteeism. According to Table 10, **there is an extensive failure to capture rotations in SA-SAMS in several provinces, but especially Limpopo.**

Table 10: % of Grade 3 learners in public schools with rotations in 2021 Term 3

	Confirmed	Suggested by regular absenteeism patterns
EC	4	36
GP	23	34
KN	7	43
LP	2	38
MP	5	30
NC	14	44
NW	5	38
Total	9	37

7 Arriving at plausible adjusted statistics on the loss of contact time

To conclude, Table 11 presents variations of the attendance percentage statistic, in line with the analysis presented above. The table refers to Grade 3 in public schools present in the 2021 Term 3 data. The 'at face value' statistic for the seven provinces is 90% (it was 91% in section 6 as there only schools which could be compared over the years are considered). This 90% figure takes into account rotation in *some* schools, but not other schools. If only regular absenteeism, and not rotation-related absenteeism, is considered then the 90% figure rises to 95%. If one takes into account the fact that 9% of Grade 3 learners are in schools where absenteeism of any kind is not reported, then *regular* absenteeism rises very slightly, and attendance declines. At the national level, the rounded attendance figure remains 95%, but for some provinces it drops by one percentage point. **While non-reporting on regular absenteeism is problematic, it makes very little difference to the aggregate statistics.** It has been assumed above that 37% of Grade 3 learners are in schools where rotations occur. If this is taken to mean that learners in these schools experience an attendance level of 50%, then **the final estimate for attendance is 78% across the seven provinces.**

Table 11: Final estimates for Grade 3

	Data 'at face value'	Regular attendance excluding effects of rotations A	% of learners in schools with no-one absent B	Attendance taking into account non-capturing of absenteeism $C = (A - B) / (100 - B) \times 100$	Estimated % of learners with rotations (from Table 10) D	Attendance taking into rotations, including non-declared $(100 - C) / 100 \times D + C / 100 \times 50$
EC	93	96	16	95	36	79
GP	82	93	1	92	34	78
KN	91	95	14	94	43	75
LP	97	98	9	98	38	79
MP	93	96	8	95	30	82
NC	82	93	1	93	44	74
NW	94	97	6	97	38	79
Total	90	95	9	95	37	78

The 78% figure is based on the NECT's finding that 35% of primary schools had rotations in the second half of 2021. Without the NECT's finding serving as an anchor, the conclusion would have been that the national attendance percentage was in the range of 72% to 86%. This corresponds to a plausible range for the indicator illustrated in

Figure 13 of 15% to 40%. **The 78% figure shown in Table 11 hides a key part of the problem, namely that losses in contact time have been *highly unequal*. If 37% of learners were in schools employing rotations, around a third of Grade 3 learners, and learners in other grades, lost around half of their contact time.**

What about the effect of rotations involving more than two groups of learners per school, meaning learners would lose two-thirds of their regular school days? The Limpopo sample discussed in section 2 had 4% of schools with three groups per school. However, the analysis of the SA-SAMS data reflected in Figure 11 above showed that less than 1% of learner-weighted schools displayed losses exceeding 50% of school days, the average for the 1% being a loss of 39%. If 1% of learner-weighted schools did have three groups per school, this would not change the 78% figure of Table 11 – the drop of 0.1 percentage points would be invisible.

8 Was attendance lower than enrolments in 2021?

8.1 The possibility of ‘ghost learners’ with 100% attendance

Up to this point, the focus has been on analysing attendance patterns among learners who attended school, especially in Term 3 of 2021. However, the attendance data analysed above also provide an opportunity to examine dropping out during the pandemic, and specifically whether the total number of learners attending school to some degree declined by half a million, as suggested by the NIDS-CRAM report discussed in section 2. This can be done by comparing *enrolments* in Term 1 of 2021 to *attending learners* in Term 3 of 2021. If the numbers match across the two datasets, then this would suggest that the pandemic did not push out an additional 500,000 learners. The earlier enrolments report¹¹, in comparing enrolments at the start of 2020, just before the pandemic began, and the start of 2021, had already found that there was no enrolment decline anywhere near the 500,000 loss seen in NIDS-CRAM. Rather, enrolments in grades R to 12 at the start of 2021 were around 50,000 lower than could be expected.

A valid concern was raised when the enrolments report was discussed. Did principals not over-report enrolments in order to protect resources flowing to the school? Teacher provisioning and school funding both rest on submitted enrolment data. Specifically, did principals keep learners from 2020 ‘on the books’ at the start of 2021, to protect resourcing? One way of examining this possibility is to compare the Term 1 enrolment and Term 3 attendance data. Attendance data are not used for resourcing purposes, raising the possibility that attendance data would be more reliable.

However, it is not impossible that schools would also have kept learners ‘on the books’ in the attendance data, marking them as present every day. The default value on SA-SAMS is that the learner attended. This must actively be switched to absent, where relevant. Schools may have kept persistently absent learners on the system, while never marking them as absent, perhaps with the expectation that they would return.

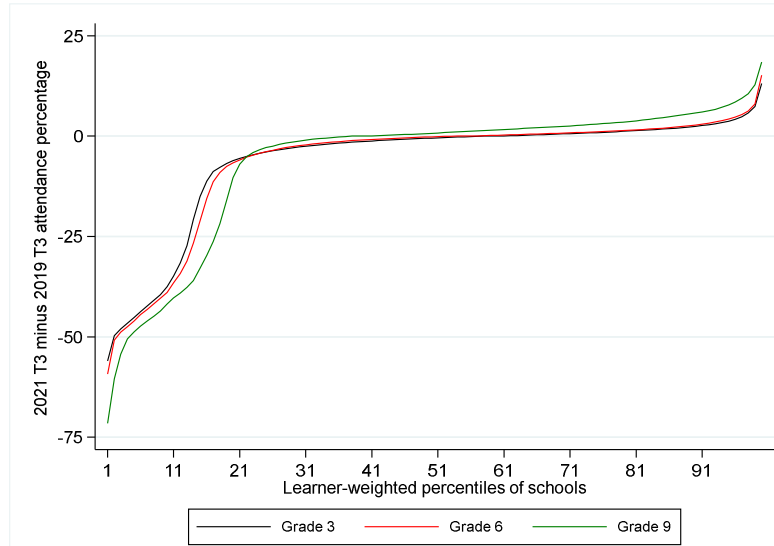
If this irregularity was the only one occurring, then one would expect the attendance percentage of schools with this irregularity to be *higher* in Term 3 of 2021, when compared to Term 3 of 2019. However, the situation is a complex one as there were many irregularities occurring all at once during the pandemic, which would also have influenced attendance percentages. Other factors that could have pushed the attendance figures *upward* in Term 3 of 2021 include the marking of learners as present on days they were not allowed to attend school, because of rotations.

¹¹ Department of Basic Education, 2021.

Figure 14 illustrates the extent to which schools had higher attendance percentages in 2021 than in 2019. This was true for 40% of learner-weighted schools with Grade 3, with 24% exceeding a gain of one percentage point or more. However, the implied number of learners is low. Assuming improvements in attendance were due only to ‘ghost’ attendees, the number of such attendees would be around 6,500, out of a total of 808,365 Grade 3 learners in the data used for the graph. This does not point to a large number of ‘ghost’ attendees.

In Grade 9, the improvement in attendance was more marked. At this level, 60% of learner-weighted schools saw an increase in their attendance percentage.

Figure 14: Schools where attendance improved at face value during the pandemic



Note: The three curves for grades 3, 6 and 9 draw from the data of 13,377, 13,086 and 5,753 schools respectively.

What was not possible to do with the available data was to examine whether the percentage of learners not absent for an entire term changed between 2019 and 2021. This is because individual learner data were only available from Term 2 of 2020. Making this comparison would have assisted in understanding the risk that there were large numbers of ‘ghost’ attendees in 2021 with 100% attendance on the system. The 2021 data point, for instance, to 22% of Grade 3 learners in public schools not being absent across terms 1 to 3 of 2021. This analysis excludes schools where no-one is ever absent. The problem is that a corresponding pre-pandemic statistic for 2019 could not be calculated with the available data.

8.2 A comparison of the 2021 enrolments and attendance data

Figure 15 below illustrates the ratio of the number of learners attending in Term 3 of 2021, even if it was just for one day, and learners enrolled in Term 1 of 2021, using the exactly the same schools for the two points in time. The comparison was conducted separately for each of the three grades. The numbers of schools available for the comparison are reported in Table 12. Comparisons were also conducted using samples of schools where the attendance percentage did *not* increase during the pandemic, according to the 2019 to 2021 comparison referred to in section 8.1. If improvements in the attendance percentage were driven by ‘ghost’ attendees, one might expect the full and sample-based comparisons to produce different results. As can be seen from Figure 15 (and Table 12), the results barely differ. This suggests ‘ghosts’ do not influence the comparison between enrolments and those attending in

2021. Such a possibility cannot be completely ruled out on the basis of these controls, but there is no indication from the data that this has been a major problem.

Figure 15: Ratio of 2021 enrolment over any attending

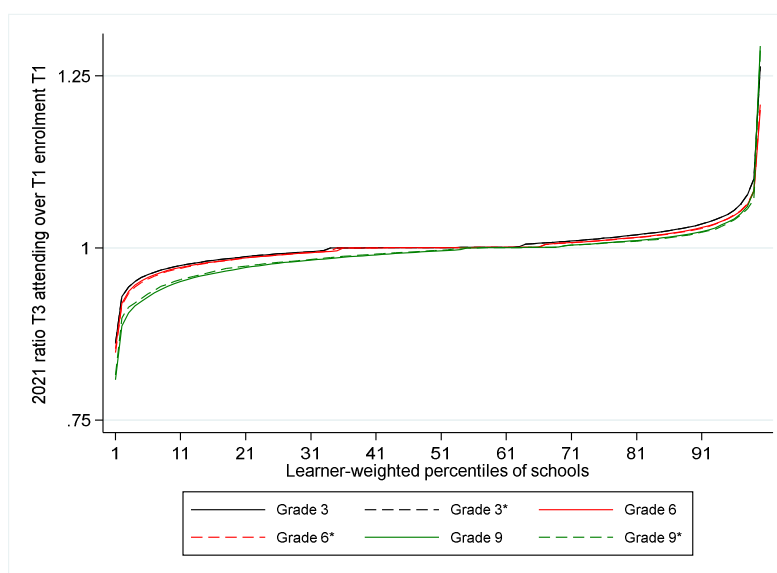


Table 12 points to learners attending being higher than enrolments in Grade 3, about the same in Grade 6, and lower in Grade 9. **There is thus no evidence that at the primary level there was major dropping out, in the sense of learners enrolling at the start of 2021, and then not attending, something that appeared to be a clear risk in the context of the pandemic. And as the earlier enrolments report has indicated, there was some, but not a large, drop in enrolments of one compares the pre-pandemic 2020 enrolment to the within-pandemic 2021 enrolment. None of this supports the NIDS-CRAM initial finding that around half a million learners dropped out due to the pandemic. Discussions with the NIDS-CRAM paper authors, and further interrogation of the NIDS-CRAM data¹², suggest that the NIDS-CRAM question, ‘Are there any learners in your household who have not yet returned to school this year’, was misinterpreted by many household respondents. Misinterpretation was made more possible by the widespread application of rotations. Many may have interpreted ‘returned to school’ as ‘returning to normal schooling’.**

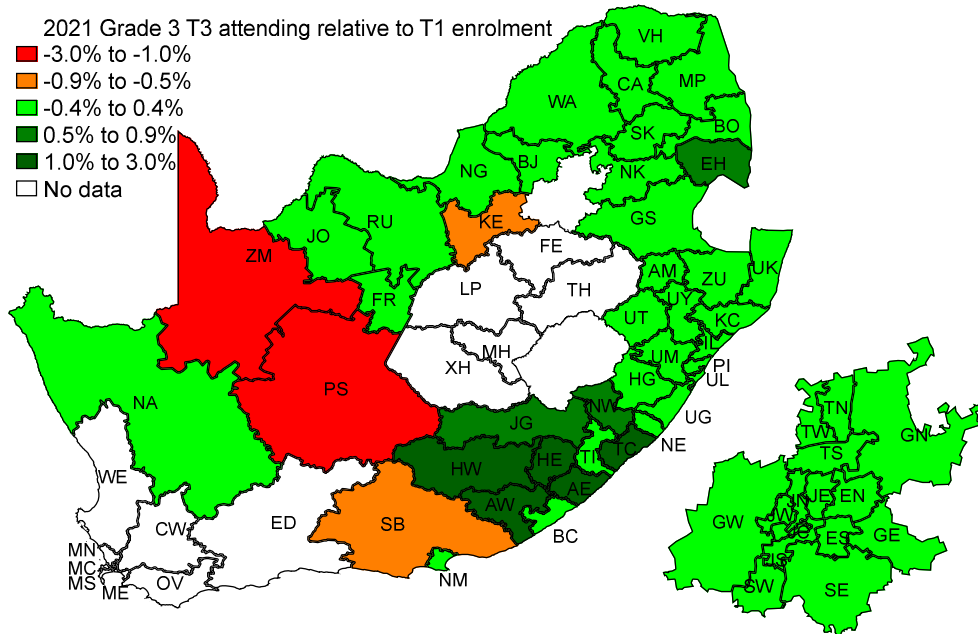
Table 12: Enrolments compared to attendance in 2021

	All public schools with data				Excluding schools with attendance % increase 2019 T3 to 2021 T3			
	Schools	T1 enrolled	T3 attending	% change	Schools	T1 enrolled	T3 attending	% change
Gr 3	10,981	689,586	700,271	1.5	6,685	398,663	405,391	1.7
Gr 6	11,099	712,704	711,939	-0.1	6,933	420,002	419,465	-0.1
Gr 9	4,947	630,427	623,456	-1.1	2,276	274,152	271,372	-1.0

Figure 16 confirms that while there may not have been an additional half a million learners dropping out as a result of the pandemic, the pandemic has affected certain education districts badly. Four districts – Sarah Baartman (SB), Pixley ka Seme (PS), ZF Mgcawu (ZM) and Dr Kenneth Kaunda (KE) – have seen attendance in Term 3 of 2021 being lower than could probably be expected.

¹² Gustafsson, 2021.

Figure 16: T1 enrolment to T3 attending change 2021



References

- Ardington, C., Wills, G. & Kotze, J. (2021). COVID-19 learning losses: Early grade reading in South Africa. *International Journal of Educational Development*, 86.
- Department of Basic Education (2021). *Impacts of the COVID-19 pandemic on school enrolments*. Pretoria.
- Gustafsson, M. (2021). *By how much has school participation declined as a result of the pandemic?* Pretoria: Department of Basic Education.
- Shepherd, D. & Mohohlwane, M. (2021). *The impact of COVID-19 in education - more than a year of disruption*. Stellenbosch: NIDS-CRAM.
- Shepherd, D., Mohohlwane, M., Taylor, S. & Kotzé, J. (2021). *Changes in education: A reflection on COVID-19 effects over a year*. Stellenbosch: NIDS-CRAM.

Annexure 1: The original data request

SPECIFICATIONS FOR THREE DATA STRUCTURES IN THREE DIFFERENT TIME PERIODS

2019 Term 2 to 2020 Term 1 – least depth of data, for pre-pandemic baseline

Fields required:

Period (year and term)*

School identifier (9-digit EMIS number)*

Grade*

Attendance (as a percentage, as in the Dashboard)

Asterisks refer to fields that uniquely identify records.

Assuming no filtering by grade, these data would have 4 periods, ±20,000 schools, and 13 grades, giving ± one million records.

2020 Term 2 to 2021 Term 2 – medium depth of data, for initial pandemic period

Fields required:

Period (year and term)*

School identifier (9-digit EMIS number)*

Grade*

Anonymised learner identifier*

Gender*

Attendance (as a percentage, as in the Dashboard)

Assuming no filtering by grade, these data would have 5 periods, and ±11 million learners, giving ± 55 million records.

2021 Term 3 – greatest depth of data, for most recent pandemic term

Fields required:

School identifier (9-digit EMIS number)*

Grade*

Anonymised learner identifier*

Gender*

Day as a date (all days for each learner, whether absent or not, included)*

Whether absent

Code indicating reason for absence (this field can perhaps be collapsed into the previous one)

Assuming no filtering by grade, these data would have ±11 million learners, and ±50 school days in the term, giving ± 550 million records.

FILTERING

The three data extracts can be filtered so that each covers only the following three grades: grades 3, 6 and 9. This is to reduce the volume of data.

STRUCTURE OF THE EXTRACT

The data can be broken up into several files, for instance according to time period, province and grade. As long as files are logically named, data can be merged programmatically during the analysis.

The format would be the regular flat text or .csv file, for instance comma-separated.