

The second main school of thought foregrounds problems of capacity - both state capacity and teacher capacity - as a key reason why there have not been more substantive improvements in learning outcomes in the poorest schools (Van der Berg et al, 2016). The most striking example of this is that 79% of Grade 6 mathematics teachers in the country cannot do Grade 6 or 7 level mathematics (Venkat and Spaul, 2015). The chapter by Taylor (N, 2019) documents the full extent of this problem, while that of Shalem & DeClerq's (2019) illustrates why this has remained an unsolved problem.

Notwithstanding the importance of either of these constraints separately, it is likely that it is the joint lack of *both* capacity *and* accountability that is the major impediment to large improvements in learning outcomes in the majority of schools (Spaul, 2015a), rather than due to just one factor. Without disagreeing with this, a number of scholars would still argue that one must begin focusing on improving capacity since capacity must precede accountability (Elmore, 2004). One can only hold people accountable for things that they can actually do. We will return to this at the end of the chapter when discussing the way forward. Despite the lack of both accountability and capacity in the South African education system over the last two decades, there have in fact been some important gains in learning outcomes across the country. Since these new results have only been available for two or three years at most (since 2016) they are not yet well known or well understood and thus deserve some attention here.

6 Trends post-apartheid

It is frequently the case that when scholars discuss the *distributions* of learning outcomes in a country that they lose sight of the levels of learning outcomes and how these are changing over time. While Van der Berg & Gustafsson (2019) in the present volume address this issue authoritatively and show that learning outcomes *have* improved in the post-apartheid period, irrespective of the data used or the grade assessed, for the general reader it is perhaps helpful to provide a stylistic overview of these trends in learning outcomes.

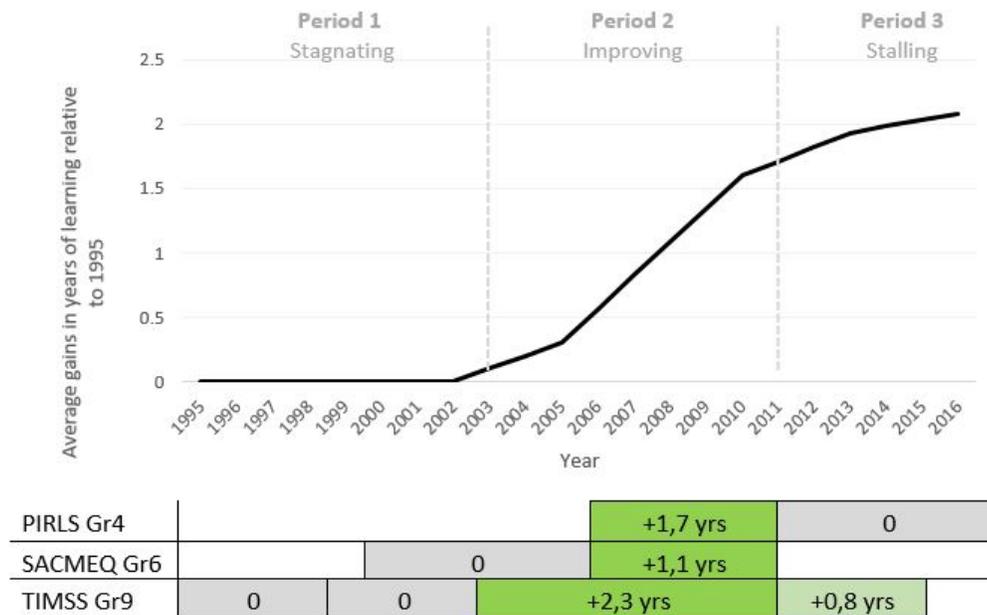
Like many middle-income countries South Africa participates in a number of international assessments every three or four years. By testing a nationally-representative sample of students from various grades and assessing a variety of subjects, these tests allow us to benchmark ourselves relative to other countries internationally and - importantly - to determine if learning outcomes are improving, stagnating or deteriorating over time. The three international assessments South Africa participates in are the Trends in International Mathematics and Science Study (TIMSS, Grade 9 Maths and Science conducted in 1995, 1999, 2003, 2011 and 2015), the Southern and Eastern African Consortium for Monitoring Educational Quality (SACMEQ, Grade 6 reading and mathematics conducted in 2000, 2007 and 2013), and the Progress in International Reading and Literacy Study (PIRLS, Grade 4; conducted in 2006, 2011 and 2016). These are the only assessments that provide psychometrically valid⁹ comparisons of learning outcomes over time.

Figure 2 below provides a stylistic overview of the trends in learning outcomes in the post-apartheid period. To do so it uses results from all waves of all three international assessments (TIMSS, SACMEQ and PIRLS) and averages the annual learning gains¹⁰ across each cycle. What is immediately clear is that there

⁹ For a discussion of comparability issues related to SACMEQ 2007 and 2013 results and for a detailed discussion of the re-scaling procedures undertaken by PIRLS to make prePIRLS and PIRLS Literacy scores comparable see Spaul & Pretorius (2019) in the current volume.

¹⁰ To be specific it calculates the difference in test scores between the start and end of the period and expresses them as a percentage of the standard deviation of the earlier period. This is converted into 'years of learning' using 0,3 standard deviations being equivalent to one year of school (see (Spaul and Kotze, 2015, p. 20)). The gray sections in the table below the graph indicate that there was no improvement in learning outcomes over that period. For SACMEQ 2007 to 2013 we use the classical

have been some gains in learning outcomes in the post-apartheid period, notably between about 2003 and 2011. There is now a growing body of reliable and consistent findings documenting these gains (see Van der Berg & Gustafsson, 2019 for a full discussion). As more corroborating evidence emerges, the position that learning outcomes have not improved in South Africa - or that “education is worse than it was under apartheid” (Ramphela, 2012) - is increasingly becoming a fringe view that is not supported by the data or serious scholars. Broadly speaking one can see three periods which could loosely be referred to as (1) a ‘stagnating’ phase (1995-2003) where learning outcomes did not improve at all (neither between TIMSS 1995, 1999 and 2003, nor between SACMEQ 2000 and 2007). (2) The ‘improving’ phase (2003-2011) where learning outcomes improved relatively quickly, supported by data from TIMSS 2003-2011, SACMEQ 2007-2013 and PIRLS 2006-2011. (3) The ‘stalling’ phase (2011-2016) where gains have flattened out as evidenced by the lower gains in TIMSS 2011-2015 and particularly the lack of any improvement between PIRLS 2011-2016.



Note: This uses comparable scores for TIMSS 1995-1999, 1999-2003, 2003-2011, 2011-2015; SACMEQ 2000-2007; PIRLS 2006-2011, 2011-2016.

Fig. 2: Gains in learning outcomes in South Africa (1995-2016) expressed as average gains in years of learning relative to 1995.

While the international assessments are the most reliable indicators of progress in education, there are also other pieces of supporting evidence, including the fourfold increase in black university graduates between 1994 (11,339 black graduates) and 2014 (48,686 black graduates) (Van Broekhuizen, 2016, p. 12), and the

test scores and standard deviation reported in Van der Berg & Gustafsson (2019) in the present volume rather than the Item Response Theory scores due to the psychometric concerns with the IRT scores discussed in Spaull & Pretorius (2019).