The South Africa PIRLS trend and the difficulties of calculating and communicating education quality trends

February 2020

Martin Gustafsson

UNIVERSITEIT STELLENBOSCH UNIVERSITY

Contents

- From Hanushek to the proficiency SDGs
- TIMSS and a bumpy SACMEQ ride
- And then an unexpected PIRLS correction
- SA policy implications: Gender etc.
- Improving our work on trends

From Hanushek to the proficiency SDGs

A CONTRIBUTION TO THE EMPIRICS OF ECONOMIC GROWTH*

N. GREGORY MANKIW
DAVID ROMER
DAVID N. WEIL

1992

A CRITICAL BREAK…

DRIVEN BY NEW DATA

Schooling, Labor-Force Quality, and the Growth of Nations

2000

By ERIC A. HANUSHEK AND DENNIS D. KIMKO*

THE ROLE OF SCHOOL IMPROVEMENT IN ECONOMIC DEVELOPMENT

2007

Eric A. Hanushek
Ludger Woessmann

DO BETTER SCHOOLS LEAD TO MORE GROWTH? COGNITIVE SKILLS, ECONOMIC OUTCOMES, AND CAUSATION

2009

Eric A. Hanushek
Ludger Woessmann

The quality vs. the quantity of schooling: What drives economic growth?

Theodore R. Breton*

Universidad EAFIT, Medellin, Antioquia, Colombia

2011

An oddity…

Human Development Report **2010**

20th Anniversary Edition

2010

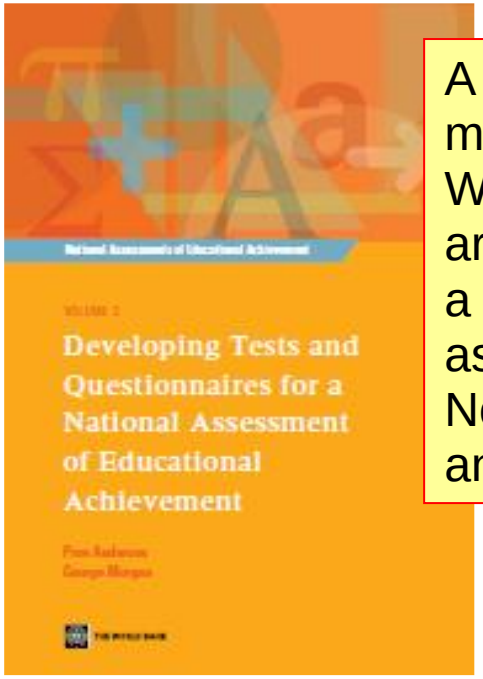
The Real Wealth of Nations:
Pathways to Human Development

STELLENBOSCH UNIVERSITY

2017

4.1.1 Proportion of children and young people:
(a) in grades 2/3; (b) at the end of primary; and
(c) at the end of lower secondary achieving at least a minimum proficiency level in (i) reading and
(ii) mathematics, by sex

Achieving a clear sense of what is required technically to gauge *changes* in proficiency has been curiously difficult.



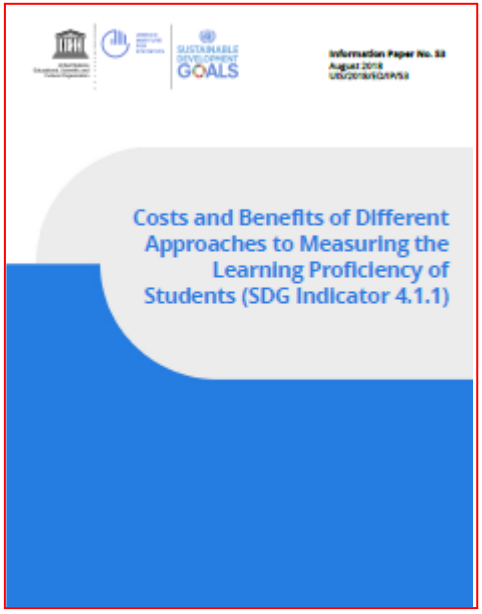
A series of manuals by the World Bank ±2008 are about running a national assessment *once.* Nothing on secure anchor items, etc!

SACMEQ

2000: Relatively good technical documentation, but nothing on IRT methods.

2007: Almost no technical documentation, when this was the first major trend analysis based on common anchor items.

2013: Again, no technical documentation.



Space and time…

The UIS has tended to prioritise comparability across countries (space), rather than within-country comparability over time. Why? Interest of donors in identifying places to target? An assumption that results barely change over time?

UNIVERSITEIT STELLENBOSCH UNIVERSITY

Historical data indicate that at least among test-takers, slow improvement more common than not. There are also ‘natural’ speed limits.

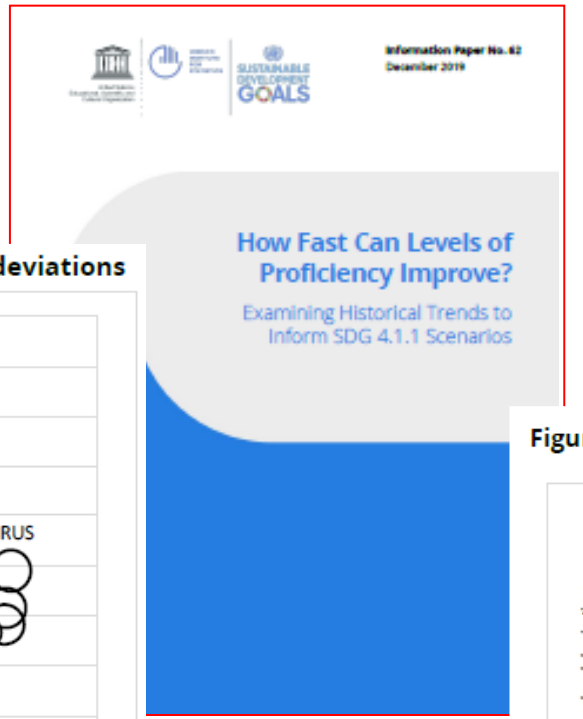
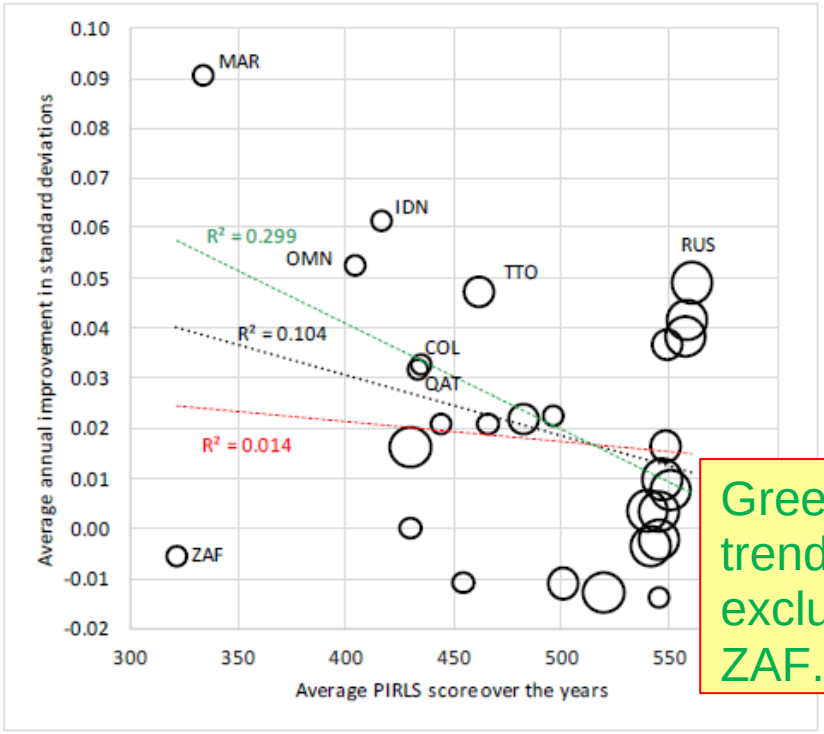
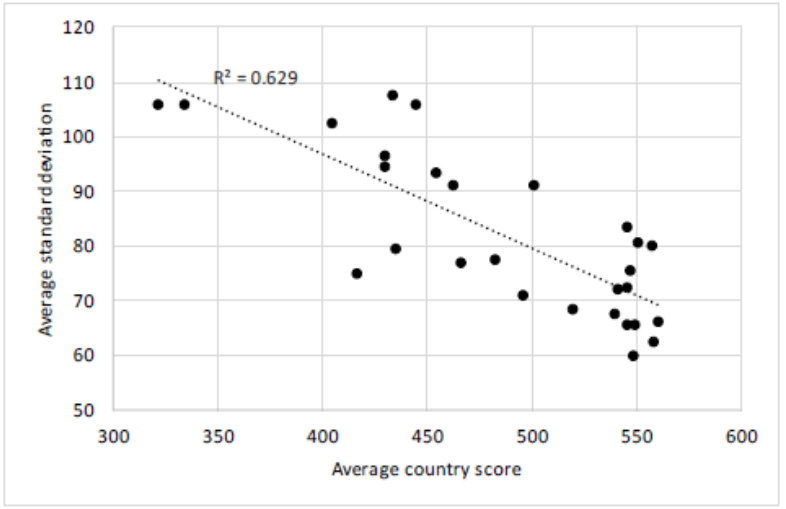


Figure 2: PIRLS improvements in terms of standard deviations



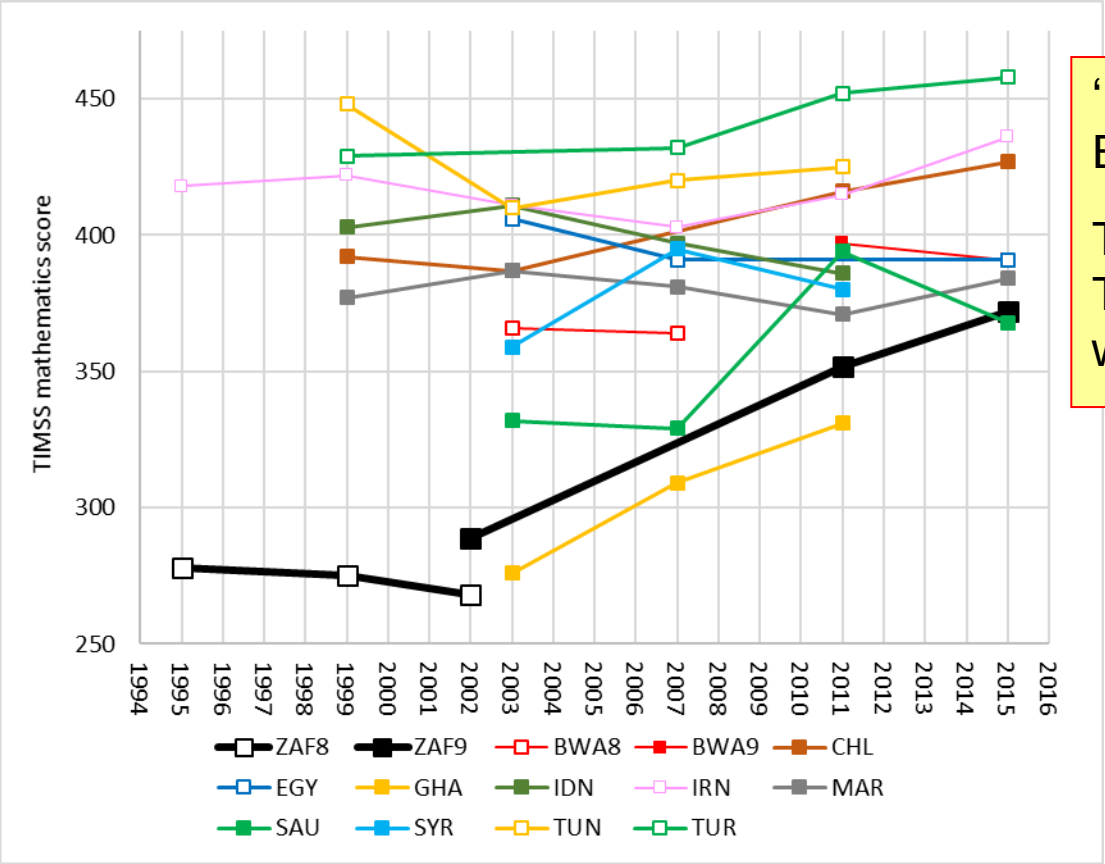
Green trendline excludes ZAF.

Figure 5: Standard deviations in PIRLS



UNIVERSITEIT STELLENBOSCH UNIVERSITY

TIMSS and a bumpy SACMEQ ride



'Becoming Botswana'…

The Nov 2012 TIMSS news was widely welcomed.

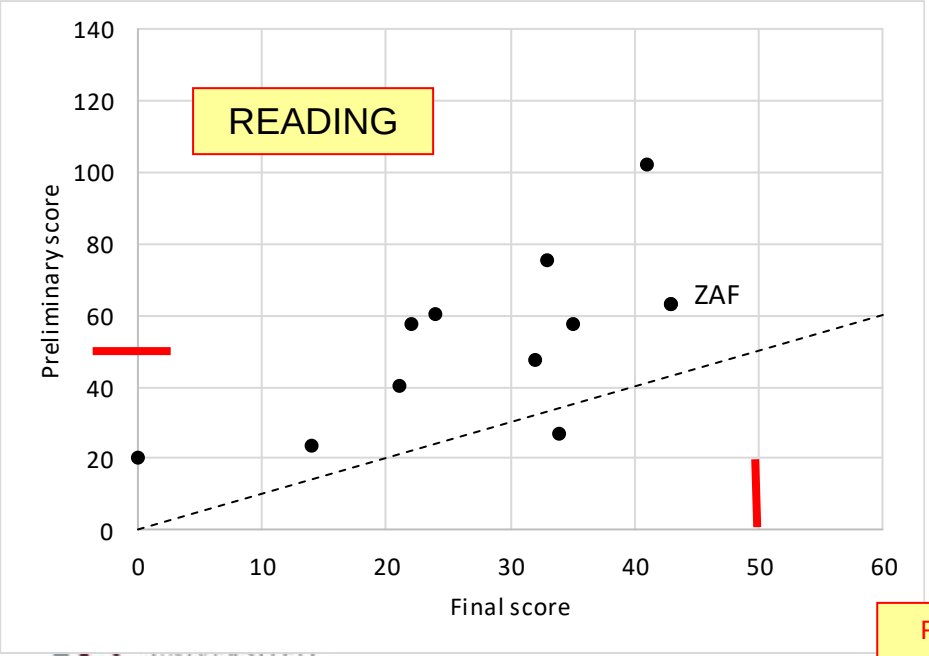
SACMEQ IV STUDY Results

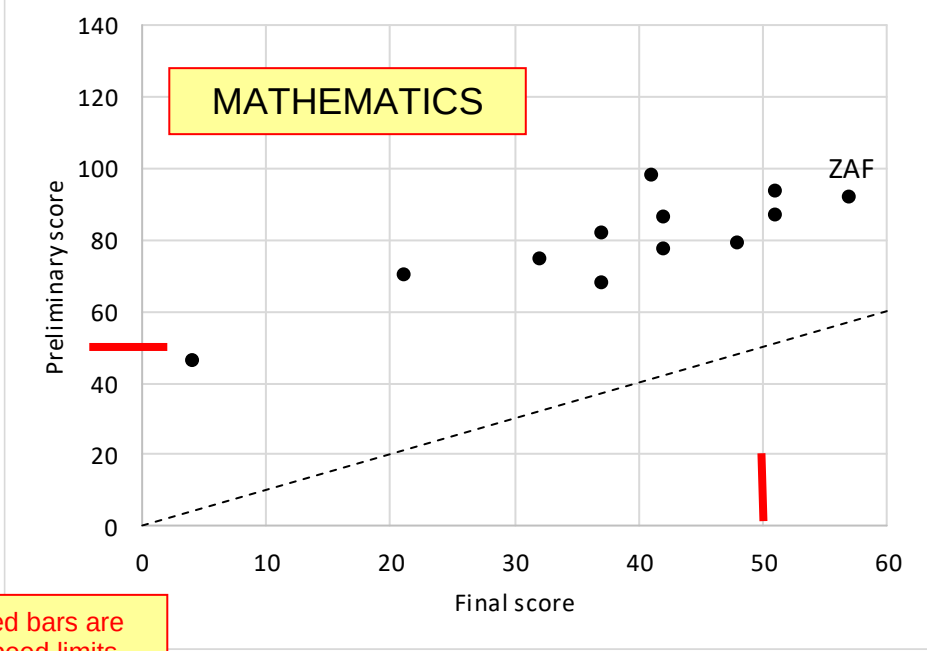
PORTFOLIO COMMITTEE
ON BASIC EDUCATION
13 SEPTEMBER 2016

THE RESULTS OF SACMEQ, TIMSS AND PIRLS 2018

ROUNDTABLE WITH DEANS OF EDUCATION AND ENGINEERING FACULTIES

https://www.education.gov.za/RoundtablewithDeansofEducationandEngineeringFaculties.aspx





Red bars are speed limits assuming 0.08 a year.

An unexpected PIRLS correction

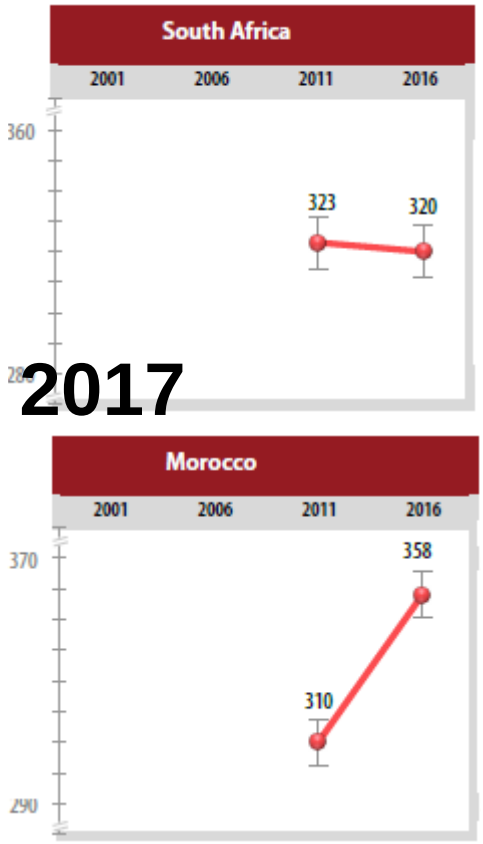
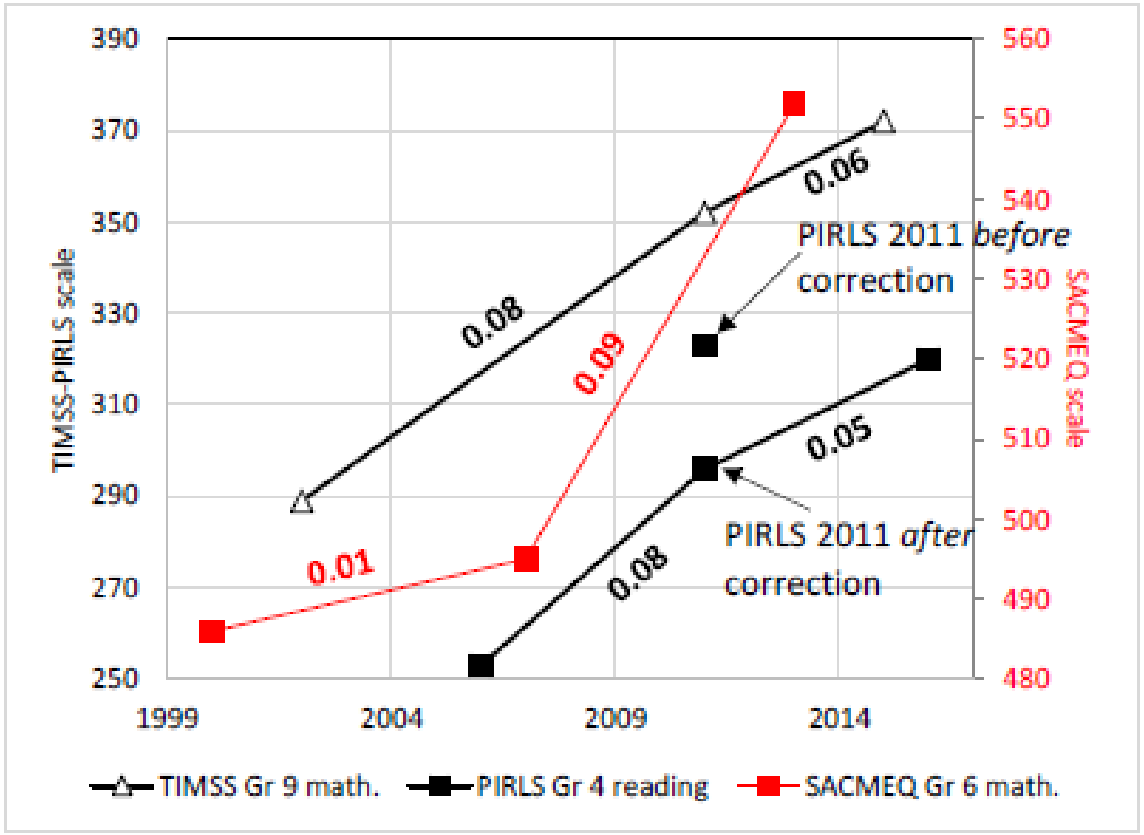


Figure 13: PIRLS, TIMSS and SACMEQ trends



Actually, moved from 295 to 320. This was the third-largest (of 43) gains for 2011 to 2016, after Morocco and Oman.

Again, few had worried that a trend exceeded a speed limit, in this case PIRLS 2006-2011.

UNIVERSITEIT STELLENBOSCH UNIVERSITY

A revised PIRLS 2011 to 2016 trend for South Africa and the importance of analysing the underlying microdata

MARTIN GUSTAFSSON

JANUARY 2020

Figure 3: Cumulative distribution for 'Baghita...' classical scores

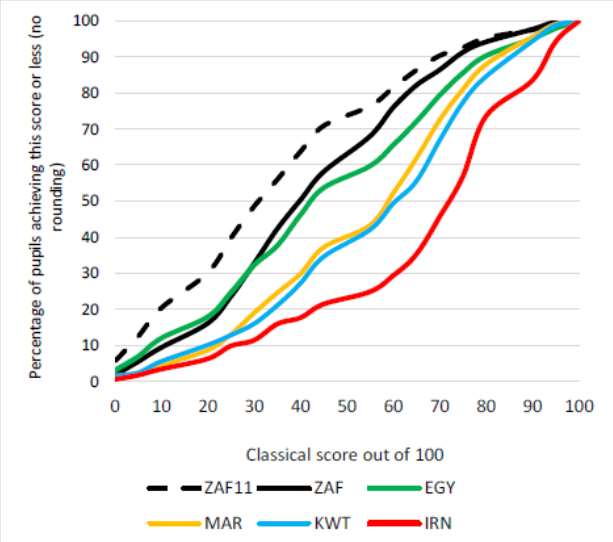


Figure 5: IRT distributions based on 51 common items

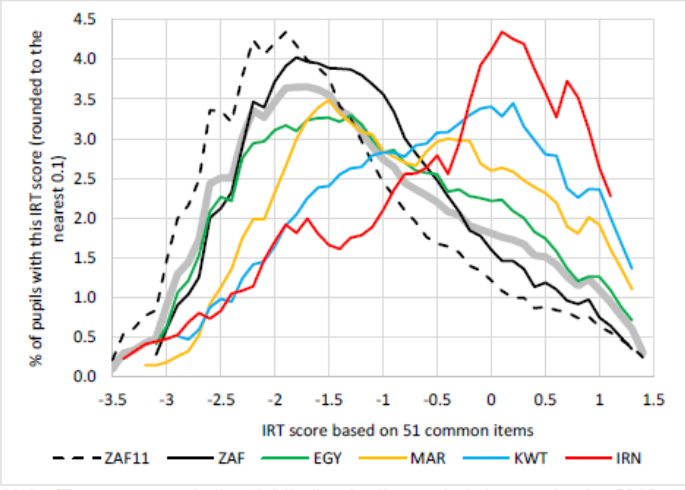


Table 1: 2006, 2011 and 2016 national Grade 4 values

PIRLS year	National Grade 4 score	Where published	Comment
2006	253	Howie *et al*, 2008: 19.	A national Grade 5 score of 302 was published in Mullis *et al* (2007: 37)
2011	323	Mullis *et* al, 2017: 29	The 2011 national score using the prePIRLS scale is 461 and appears in Mullis *et al* (2012: 39).
2016	320	Mullis *et* al, 2017: 29	

Figure 4: 2011-2016 gains by item

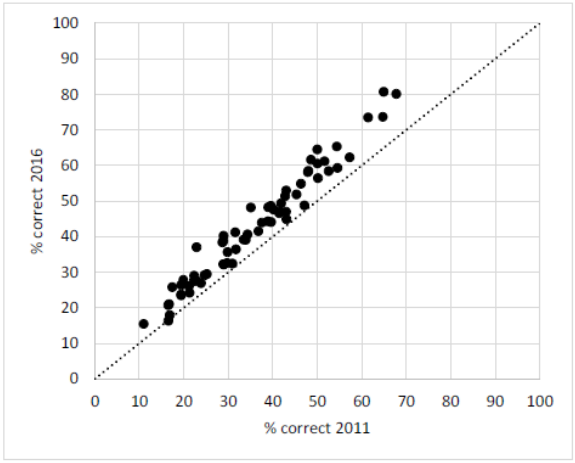
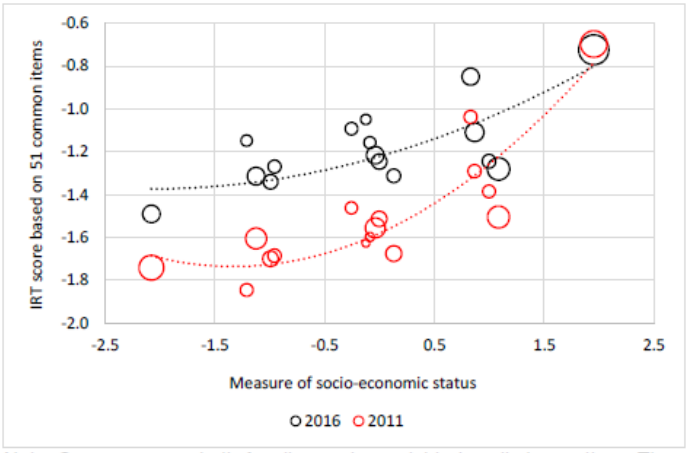
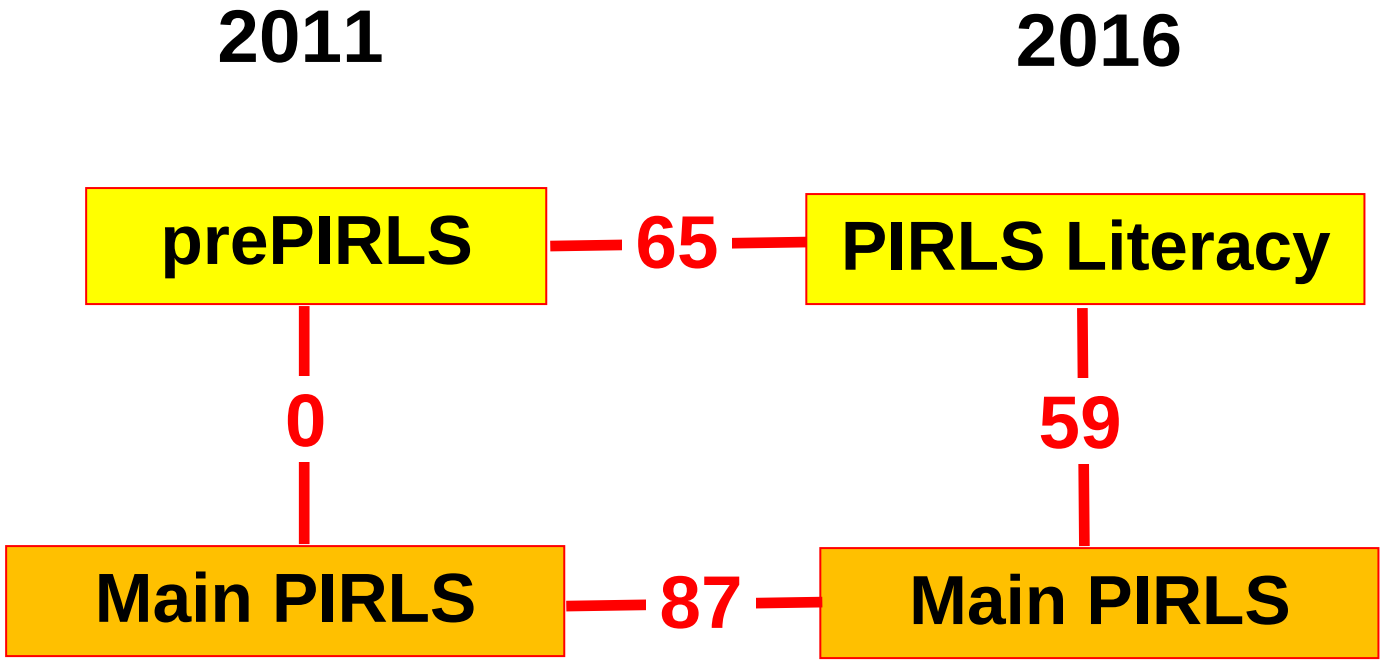


Figure 12: Percentile mean IRT scores in South Africa's 2016 PIRLS Literacy



Colombia took both in 2011. This is what Boston College used for the South Africa conversion. Not clear why.



SA policy implications: Gender etc.

- TIMSS Grade 9: Between 2002 and 2015, from pro-boy to pro-girl. Had boy progress been applicable to everyone, overall gain would have been 7% smaller.
- In PIRLS 2016, South Africa displays second-largest pro-girl advantage (of 50), after Saudi Arabia (!).
- Non-sector factors to ‘protect’: definitely urbanisation, highest education in household; maybe less poverty.
- Sectoral factors: access to books, curriculum clarity, subject knowledge of younger teachers.

Improving our work on trends

1. Develop capacity in psychometrics, inside and outside government.
2. Insist on a more rigorous approach to performance trends on the part of global bodies: UIS, World Bank, WEF.
3. Insist on transparent technical documentation (SACMEQ, but also national assessments) and have zero tolerance for isomorphic mimicry.
4. Make microdata publicly available (á la TIMSS/PIRLS).
5. Separate players (ministries) from referees (assessment authorities).

The case for statecraft in education: The NDP, a recent book on governance, and the New Public Management inheritance

MARTIN GUSTAFSSON

UNIVERSITEIT STELLENBOSCH UNIVERSITY