Assessing learning proficiency levels and trends for Sustainable Development Goal 4.1
A focus on Africa

January 2022
Martin Gustafsson
University of Stellenbosch
Main findings
Aim is to examine the availability and reliability of statistics on SDG 4.1 in Africa

- Availability of official SDG 4.1 statistics about as low in Africa as elsewhere, but at primary level relatively good in Africa.
- Population-weighted proficiency statistic for lower primary in Africa: 29% (57% in rest of developing world).
- Confirmed, the higher the grade the lower (worse) the statistic.
- For 2014 to 2019, 14 of 55 countries have a trend – largely thanks to PASEC.
- Average annual gain is 4.3 proficiency percentage points a year, 3.0 in rest of developing world. Yet ‘speed limit’ is around 3.0.
Main findings (contd.)

- By July 2021 **69%** of a year lost by Africa’s learners due to the pandemic.
- Children in countries with relatively good assessments:
  - **50%** in countries with international assessment in last 10 years.
  - Figure rises to **78%** if sample-based national included.
- **Internal SGD 4.1.1 quality checks:**
  - Ranking consistency across: (a) time; (b) level; (c) subject.
- **Quality checks with a broader range of data:**
  - Consistency over time in terms of background variables.
  - Across-programme comparisons.
Recommendations

1. Focus on improving measurement using indicators already established and approved.
2. Focus on improving PASEC and SACMEQ, which are a vital resource not just for 28 participating countries.
   - Assessment design and analytical capacity at the national and continental level.
   - Accept that the two programmes are young and must undergo design improvements.
3. Improve national ownership of proficiency statistics, including quality control and selection of the right statistics.
4. Promote holistic data analysis, comparing (triangulating) data from different assessments, background questionnaires, household surveys, different countries.
Complex institutional backdrop

- Global monitoring bodies have often focussed especially on fine-tuning *across-country comparisons*, while countries pay more attention to *national trends over time*.
- **Sample-based assessments** clearly more reliable, but country investments in systems and debates around improvements and often revolve around examinations (or censal assessment). Bringing these two worlds within the same orbit would be good.
- **Quality control**, by UIS and countries, made difficult by political context. More *generic*, as opposed to *country-specific*, evaluations of statistics is one route UIS can follow.
Consistencies and inconsistencies

Of 47 2014-2019 trends, 42 are from PASEC, and of these 14 are at the 3.0 ‘speed limit’ or above.
Consistencies and inconsistencies (contd.)

PASEC gains all the more remarkable against the background of large enrolment increases. The latter probably driven more by population growth than improved participation rates.
Consistencies and inconsistencies (contd.)

This raises questions around the comparability of the PASEC samples.
Consistencies and inconsistencies (contd.)

Assuming a clean measurement process, Burundi stands out for (1) performing exceptionally well in Grade 6 mathematics relative to reading, and (2) having an unusual curriculum focus in mathematics.

Background work not included in the published report.
Consistencies and inconsistencies (contd.)

Burundi’s patterns are not the only ones in PASEC that raise questions. Moreover, the evaluation covered only within-2014 consistencies as details on 2014-2019 anchor items were received later.
Consistencies and inconsistencies (contd.)

Testing via households helps coverage of the out-of-school, but the disadvantage is large selection effects, which worsen the poorer the household.
Consistencies and inconsistencies (contd.)

Table 7: Countries with large ranking discrepancies in two-way comparisons

<table>
<thead>
<tr>
<th>S EPR</th>
<th>S LPM</th>
<th>WP</th>
<th>HTS</th>
<th>YOUTH</th>
</tr>
</thead>
<tbody>
<tr>
<td>S LPR</td>
<td>BDI</td>
<td>BDI</td>
<td>NER</td>
<td>GHA ZAF ZWE</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NER</td>
</tr>
<tr>
<td>S EPR</td>
<td>BDI</td>
<td>BDI</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S LPM</td>
<td>LSO</td>
<td>COD</td>
<td>NER</td>
<td>GHA LSO ZAF ZWE</td>
</tr>
<tr>
<td></td>
<td>COD</td>
<td></td>
<td></td>
<td>NER TCD</td>
</tr>
<tr>
<td>S EPM</td>
<td>ETH</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WP</td>
<td>BDI</td>
<td></td>
<td></td>
<td>COD</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BEN</td>
</tr>
<tr>
<td>HTS</td>
<td></td>
<td></td>
<td></td>
<td>COD DZA EGY GHA MAR NGA ZAF ZMB</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BEN BFA CAF GIN SEN</td>
</tr>
</tbody>
</table>

World Bank harmonised test score (HTS), behind the Human Capital Index, produces rankings which are inconsistent with other programmes.
**Consistencies and inconsistencies (contd.)**

**Table 8: Comparison of proficiency levels across series in Africa**

<table>
<thead>
<tr>
<th></th>
<th>S LPM</th>
<th>S EPR</th>
<th>S EPM</th>
<th>S LSR</th>
<th>S LSM</th>
<th>WP</th>
<th>MICS R</th>
<th>MICS M</th>
<th>YOUTH</th>
</tr>
</thead>
<tbody>
<tr>
<td>S LPM</td>
<td>11</td>
<td>-13</td>
<td>-45</td>
<td>-26</td>
<td>-14</td>
<td>-18</td>
<td>-25</td>
<td>-26</td>
<td>-19</td>
</tr>
<tr>
<td>S EPR</td>
<td>-13</td>
<td>-33</td>
<td>-11</td>
<td>-14</td>
<td>-18</td>
<td>-14</td>
<td>-25</td>
<td>-19</td>
<td>-6</td>
</tr>
<tr>
<td>S EPM</td>
<td>-26</td>
<td>-45</td>
<td>-11</td>
<td>-14</td>
<td>-18</td>
<td>-14</td>
<td>-25</td>
<td>-19</td>
<td>-6</td>
</tr>
<tr>
<td>S LSR</td>
<td>-14</td>
<td>-21</td>
<td>-18</td>
<td>-19</td>
<td>-3</td>
<td>-1</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>S LSM</td>
<td>-18</td>
<td>-25</td>
<td>-26</td>
<td>-19</td>
<td>-6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WP</td>
<td>-18</td>
<td>-38</td>
<td>-8</td>
<td>3</td>
<td>-1</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MICS R</td>
<td>-9</td>
<td>-16</td>
<td>-5</td>
<td>-1</td>
<td>19</td>
<td>22</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MICS M</td>
<td>-16</td>
<td>-20</td>
<td>-7</td>
<td>-4</td>
<td>-13</td>
<td>-10</td>
<td>-8</td>
<td>-6</td>
<td></td>
</tr>
<tr>
<td>YOUTH</td>
<td>42</td>
<td>31</td>
<td>43</td>
<td>54</td>
<td>64</td>
<td>69</td>
<td>53</td>
<td>62</td>
<td>68</td>
</tr>
</tbody>
</table>
**SDG official vs MILO**

Large marker is most recent SDG 4.1.1b official, or the 2019 PASEC figure, if different.

Small marker is MILO pre-pandemic (with-pandemic in the case of Kenya reading).
Conclusion

Careful evaluation of the valuable and important data and statistics we have on learning is necessary, by national analysts and those looking at regional and global trends. Not doing this exposes policymaking to serious risks. But doing this well requires a stronger focus on building analytical capacity, especially at the national level.