What have we learned about the use of learning assessment data? An overview of recent research and trends

As a subset of learning assessment systems, large-scale learning assessments (LSLAs) can be defined as "system-level assessments that provide a snapshot of learning achievement for a given group of learners (based on age or grade) in a given year and in a limited number of domains" (UNESCO, 2019: 17).

2019 has been a productive year for research on LSLAs, their use and potential impact on education policy. While the use of LSLAs to measure learning is still a relatively new phenomenon (Fischmann et al., 2018), an increasing number of countries now take part in regional and/or international measurements. At an international level, LSLAs involve either multiple countries from different regions (such as PISA, administered by the OECD) or a group of countries from a single region (such as ERCE, organized by UNESCO in Latin America). Such assessments are usually curriculum-based, designed to measure how well students have acquired the curriculum, or skills-based, to measure how students apply the knowledge and skills they have acquired (Addey and Sellar, 2019). The scope of such tests traditionally focussed on literacy and numeracy and now includes domains such as digital skills, socio-emotional skills, and civics and citizenship knowledge (UNESCO, 2019).

Within the context of the Sustainable Development Goals (SDGs), LSLAs have become an important tool for measuring the quality of education and for understanding gaps at national and global levels. The 2019 edition of the SDG 4 Data Digest produced by the UNESCO Institute for Statistics provides the latest developments and strategies that countries can adopt to produce data for reporting on the different international SDG4 indicators as well as their national priorities (UIS, 2019). While LSLAs may have the potential to influence education policy positively, there are challenges to consider with regard to their design and

implementation, as well as the use of data obtained through them.

The motivations behind participation

Addey and Sellar point to a number of reasons why countries participate in LSLA's. These include: a need to measure educational outcomes of implemented policies and identify other policies that "work" in order to drive further policy changes, to show belonging to a group of countries with shared values and goals, to advance or change domestic policies, to obtain international funding, and to improve policy and pursue economic goals (Addey and Sellar, 2019).

UNESCO lists the following benefits in participating in LSLA's:

- Providing access to networks at regional and international levels, both for researchers and policymakers;
- Provoking dialogue among different groups and allowing for catalyzing debate on education;
- Producing evidence for educational phenomena and advancing the use of technical and complex education indicators;
- Motivating regulatory and behavioural policy reforms, including on teaching, learning and assessment;
- Attracting media attention and increasing transparency regarding education system outcomes and human capital development in national and cross-national contexts; and
- Developing capacities of professionals who participate in the assessment – from a technical and operational perspective. (UNESCO, 2019: 23)

Addey and Sellar also discuss possible reasons for not participating. These may be ideological (as in "making a statement"), lack of resources and capacity, concerns about performance, or simply because LSLAs are not a policy instrument considered as relevant in a given context (Addey and Sellar, 2019:10).

Research on whether LSLA's have an impact on educational policies is so far inconclusive (Addey and Sellar, 2019). As Fischmann et al. suggest, "some countries have the same ILSA results but implement different policies, and vice versa, even though we cannot definitively say that these policy convergences/reactions can be explicitly linked to ILSAs participation and performance" (Fischmann et al., 2018: 18).

Understanding limitations and risks

UNESCO highlights four issues within the *design* of LSLAs that could be negatively affect their usefulness. First, the conceptualization of what is valuable in learning could be severely constrained due to the limited range of competencies that LSLAs measure. Second, learning outcomes measured by LSLAs are one among many indicators of education quality. Third, LSLAs risk restricting what is valued in education due to the insufficient attention to the breadth of knowledge and skills they encompass. Finally, "the inclusiveness imperative" is not systematically present in the design and implementation of LSLAs, which could increase the exclusion of vulnerable and marginalized students (UNESCO, 2019: 12).

The limitations associated with LSLAs also relate to the use of their results. Limited use of assessment data could stem from a lack of institutional support to assessment practices, and scarce resources to fund, design, implement and disseminate LSLAs. Conversely, an excessive emphasis on assessment results could have a negative effect on policymaking – for instance, if the focus shifts from substantial reforms to improving the test scores. Finally, the combination with accountability schemes could have unintended consequences, often reflected into the adoption of "teaching to the test" dynamics (UNESCO, 2019: 13).

Based on the available literature, Raudonyte identifies a series of factors influencing the use of learning data in education policymaking, grouped into four categories: reliability and relevance of the information provided, financial and technical capacities, coordination and dissemination channels, and political and institutional factors (Raudonyte, 2019: 9). Reliability and relevance of learning data are key for policymaking, and if data are considered flawed or irrelevant, they might be disregarded altogether.

Insufficient technical and financial resources can be an obstacle for the implementation of LSLAs and the analysis of the data for policy decisions. Ineffective coordination often hinders the use of assessment results, so the involvement of stakeholders in the design and data analysis of LSLAs is key. This can consequently improve the quality of the information produced. Finally, political factors such as a fear of disappointing findings from LSLAs, that might make leaders unpopular, could also translate into underuse of assessment data (Raudonyte, 2019: 20-23). In addition, decision-making based on data from LSLAs could be misguided if, for instance, a single test score is used without considering other qualitative or quantitative data sources, or when data is instrumental to legitimize pre-defined government policies (Raudonyte, 2019).

Wyatt-Smith et al. examine the emergence of big data and digital learning assessments, and their influence on education policy and school practices. While digital learning assessments are "the conversion of standard testing practices to an online form" (Wyatt-Smith et al., 2019: 2), their evolution has moved to a more complex Computer Adaptive Testing (CAT) that is "designed to adapt the test to the test taker's ability to answer questions, thus transforming the nature of the tests, the test taker's experience of the test, and potential modes of test data analyses, log-file data, and feedback" (Wyatt-Smith et al., 2019: 5). CATs have the potential to provide faster feedback to teachers and students, but also present challenges. For instance, the increasing support from for-profit edu-businesses and philanthropic organisations poses the question about who does what to define the purposes and functioning in our educational systems (Wyatt-Smith et al., 2019: 7). A bigger issue is the "de-professionalisation" of teachers due to digital disruption in schooling. The authors argue that teacher professional judgment should be central to education quality and used alongside the information produced through digital learning assessments and that teachers should be given the opportunity to "reprofessionalise" to develop the skills required to turn data into pedagogical action (Wyatt-Smith et al., 2019).

Regional trends

In Latin America, Bruns et al. found that "the region is unique among developing regions in the high number of countries that can report globally-benchmarked learning results at all three measurement points recommended for monitoring SDG 4" (Bruns et al, 2019: 12). The region appears to have followed a coordinated effort to implement LSLAs at a national, regional, and international level. The authors suggest that the experience with the measurement of learning outcomes in Latin America over the last 20 years has valuable lessons for sub-Saharan Africa, where currently learning data remain limited, in particular regarding regional efforts to implement a common LSLA (Bruns et al., 2019: 54).

Measuring 21st-century skills

SDG 4.7 reporting requires the measurement of 21st-century skills or transversal competencies. In a study on eight Asian countries, Care, Vista, and Kim found that while measurement of transversal competencies is featured in educational policies in some countries/jurisdictions in the region, there are challenges in terms of implementation at the school level. These include: "lack of teacher professional development, clear guidance and guidelines, as well as lack of

support in terms of access to assessment tools" (Care, Vista and Kim, 2019: xi). The authors propose four "big issues" for reflection. Firstly, the need for a better understanding of the nature of transversal competencies or 21st-century skills. Secondly, the ways in which the social skills of communication and collaboration might be advanced and assessed. Thirdly, the likelihood of some school subjects lending themselves more easily to the teaching, learning, and assessment of particular TVC than others. Finally, the recognition of the capacity of current assessment tasks to encompass learning beyond the traditional academic requirements (Care et al., 2019: 32).

There is a growing interest and agreement on the usefulness of data on learning outcomes to improve national education policy and to assess the progress toward the global education goals. National and cross-national learning assessments have the potential to drive educational progress by providing timely and quality information to policymakers (UNESCO, 2019), but their use is not exempt from challenges and risks, and questions remaining on how the data obtained through LSLAs is used.

References

1. Addey, C.; Sellar, S. 2019. Is it worth it? Rationales for (non)participation in international large-scale learning assessments. Education, research and foresight: working papers, 24. Paris: UNESCO.
2. Bruns, B.; Akmal, M.; Birdsall, N. 2019. The Political economy of testing in Latin America and Sub-Saharan Africa. CGD Working Paper 515. Washington, DC: Center for Global Development.
3. Care, E.; Vista, A.; Kim, H. 2019. Assessment of transversal competencies: current tools in the Asian region. Bangkok: UNESCO Bangkok.
4. Fischman, G. E.; Marcetti Topper, A.; Silova, I.; Goebel, J.; Holloway, J. L. 2018. Examining the influence of international large-scale assessments on national education policies, Journal of Education Policy, 34:4, 470-499.
5. Raudonyte, I. 2019. Use of learning assessment data in education policy-making. Paris: UNESCO-IIEP.
6. UIS. 2019. SDG 4 data digest: How to produce and use the global and thematic education indicators. Montreal: UIS.
7. UNESCO. 2019. The promise of large-scale learning assessments: acknowledging limits to unlock opportunities. Paris: UNESCO.
8. Wyatt-Smith, C.; Lingard, B.; Heck, E. 2019. Digital learning assessments and big data: Implications for teacher professionalism. Education, research and foresight: working papers, 25. Paris: UNESCO.

Bookmark this

[View PDF](#)