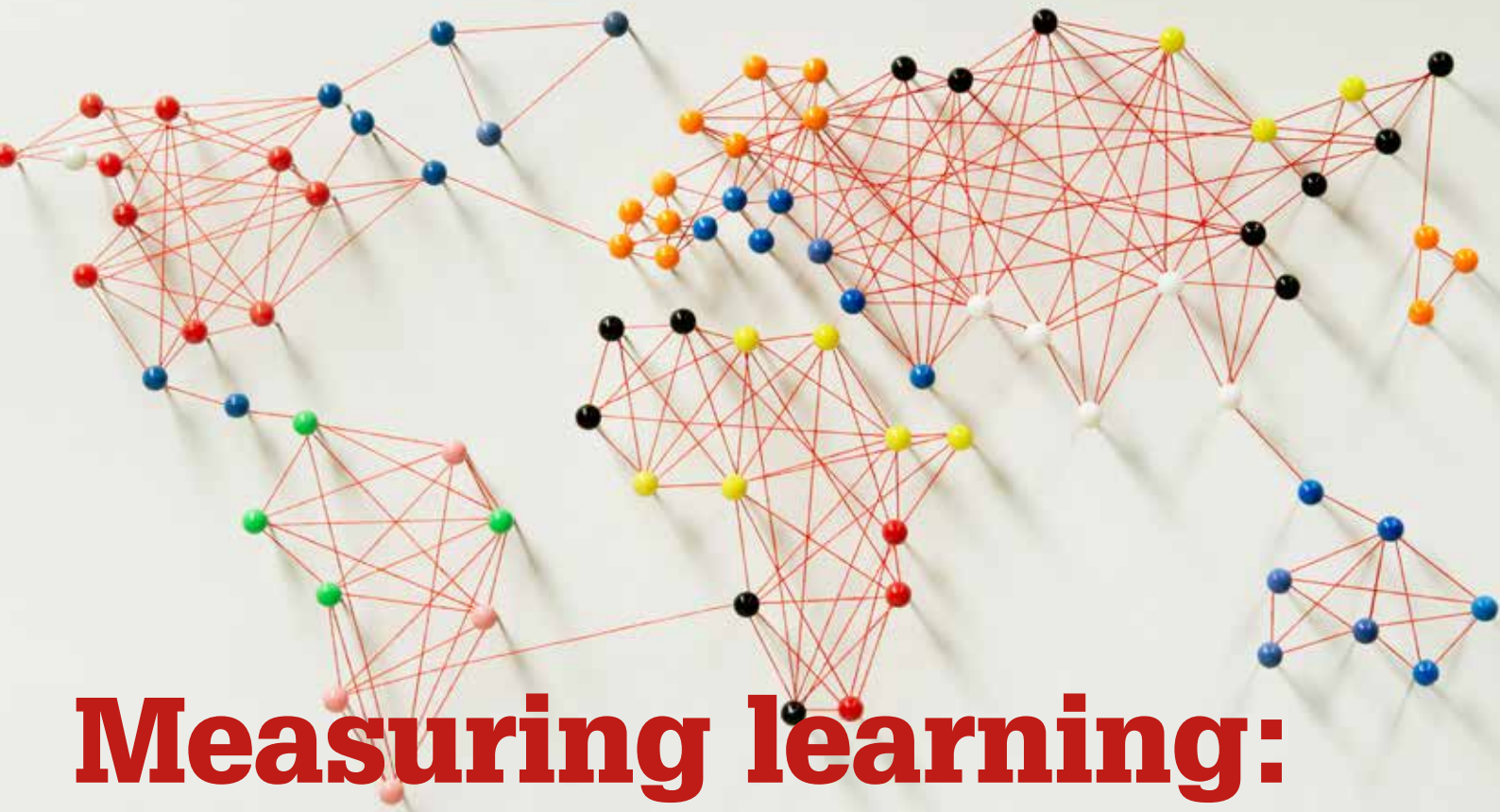

Measuring learning:
Why we need to work together

One of the keys to increasing the quality and comparability of learning measures is cooperation within the international community to develop common benchmarks to measure learning outcomes, argues **Jeaniene Spink.**

Jeaniene Spink is the Research Director for ACER's Educational and Development Research program.

A cooperative effort is vital, since comparable measures of learning outcomes across locations, population groups and time do not exist in many parts of the world.

The good news is that the new Sustainable Development Goals (SDGs) have placed quality at the heart of the global education development agenda for the next 15 years. The bad news is that we will be unable to determine the success of the SDGs in 2030 if we fail to develop a common understanding of what we mean by 'quality' and with no defined process for developing benchmarks of performance against a common measure.

As Rachel Outhred explained in 2015 in a post for the Economist Intelligence Unit's *Insights* magazine, developing comparable measures of learning outcomes is a necessary first step in the global discussion on learning improvements. While that's surely a view that is wholly endorsed by many in the education development sector, the crux of the issue is that comparable measures can only be possible through the development of a common set of universal metrics.

So what is a metric, or set of metrics? A metric, or scale, is not a test; it is like the lines on a ruler or other measuring device. In the same way as we use a ruler to measure length and a thermometer to measure temperature, placing the length and the temperature respectively on an agreed measurement scale, a universal metric (for reading, for example) is a measurement scale on which we can locate student reading skills. It clearly describes a successive set of skills that students come to possess as they acquire and develop reading proficiency. In order for teachers and policymakers to know where students lie on these metrics, a multitude of assessments or tests can be developed to assess their skills. Similarly, teachers and policymakers may identify locations on the metric that can act as benchmarks

to define, say, 'minimum standards' or 'basic literacy'.

It is true that a multitude of assessments are currently being implemented across the developing world. A joint study by UNICEF and the Australian Council for Educational Research (ACER) in Eastern and Southern Africa found there were 58 student assessments carried out in 21 countries in the region, yet ascertaining trends and identifying common areas and strategies for development investment as a result of these assessments was not possible. Because each of these 58 assessments used its own measurement scale or metric, comparing the results was not possible. It was like having 58 sets of bathroom scales, all calibrated in different units. Even what it meant to have 'limited learning outcomes' was defined differently by the 58 assessments.

At present there is no common reference point that allows student learning to be measured across contexts and over time, and we have no common benchmarks to define what we mean by 'limited learning outcomes'. Without this basic starting point, the risk is that scarce resources in the developing world may not be effectively targeted in pursuit of meaningful and lasting improvements in student learning by 2030.

The development of universal metrics as proposed by the UNESCO Institute for Statistics has the potential to revolutionise investment strategies in education. Providing a common metric for measuring student learning will not only impact student assessment, but also teacher development strategies, curriculum reform platforms and national education standards setting. A universal metric neither precludes the use of multiple assessments nor replaces existing student assessments. A universal metric instead provides a common reference point: in other words, a global measurement scale.

But how do we develop this universal common metric? The most cost effective way is by building on existing assessments; that is, by using existing test items from multiple assessments implemented in a range of educational settings across the world so that a conceptual measure of student learning can be built. This means that a universal scale will be exactly that – universal. It will not be based on a single assessment program or instrument, but will be an amalgamation of all existing measures of student learning.

Some critics have argued that a common universal metric is not possible, especially in the context of reading, given the cultural and language-specific elements of reading literacy. This assertion, however, is challenged by the international scale construction that has successfully underpinned such assessments as the Programme for International Student Assessment and the Progress in International Reading Literacy Study. More recent research conducted by ACER has drawn on data from thousands of student responses to more than 500 items and on the outcomes of 26 000 pair-wise comparisons of items from a multitude of different assessments, to support the drafting of reading metrics. The research indicates that reading metrics that begin with basic reading prerequisite skills, that often vary across languages, and progress to include sophisticated information retrieval and reflection upon text can be built. As Ms Outhred correctly points out, while such a metric inevitably involves compromise, and may not be perfect, there is strong research evidence to show that a robust scale fit for the purpose can be achieved.

ACER has for many decades designed and run various assessment programs around the world, and provided support for a variety of international, regional and national education programs. Through the Centre for Global Education Monitoring, ACER is using its long experience and established expertise to advance models of good practice related to different aspects of education, particularly in developing countries. A recent memorandum of understanding between ACER and the UNESCO Institute for Statistics means two of the world's leading educational research centres are now collaborating to support standardised and definitive reporting of countries' progress towards achieving learning for all.

The development of a set of common universal metrics for reading and mathematics may not be easy, but the fact is the global community has signed up to the challenges of the SDGs. We therefore have a responsibility to build what is likely to become one of the most important drivers of quality improvement for the lives of millions of children across the world. The time to act on this is now. ■

In this issue

In this issue of *International Developments* we look at the purposeful collection of educational data through progressive achievement testing to enable teachers to establish where students are in their long-term learning, diagnose individual strengths and weaknesses, identify the best next steps for action, decide on appropriate evidence-based interventions, monitor the progress students make over time, and evaluate the effectiveness of their own teaching decisions and approaches.

We also explore how a new primary years assessment is helping teachers, curriculum designers and policymakers to better measure the learning achievement of students in South East Asia; investigate the impact of a professional learning program for aspiring school principals in Indonesia; and find out how a national assessment program to support teaching and learning in Saudi Arabia is also developing organisational capacity.

On the policy front, we report on ACER's analysis of the impact of large-scale assessments on education policy in the Asia-Pacific region that is helping stakeholders improve the design and usefulness of assessments, and find out why the citizen-led approach to the collection of information about schooling and children's learning is improving educational monitoring and informing policy making in India, Mali, Senegal, Kenya, Tanzania and Uganda.

Further information

A version of this article was first published in the Economist Intelligence Unit's *Perspectives* online magazine. www.eiuperspectives.economist.com