Learning progressions as an inclusive solution to global education monitoring

Ray Adams, Jen Jackson, and Ross Turner, on behalf of the Australian Council for Educational Research Centre for Global Education Monitoring (ACER-GEM).

The Sustainable Development Goals (SDGs) mark a welcome shift in global discussion of education, because now we are talking not only about getting children into school, but also about making sure that they are learning. At the same time, the SDGs introduce a new layer of complexity, because we need to define what quality education means in a way that is meaningful across international contexts. Indicator 4.1.1 is an example:

> 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.

This indicator has a number of terms that are unlikely to have consistent meanings across all the education systems that will participate in SDG reporting. What does "grades 2/3" mean in an African country compared to a European country, or to the United States? What does "minimum proficiency" mean in a high-income country, compared to a country that is economically developing? Are "reading" and "mathematics" the same around the world? Unless we can agree on answers to these questions, Indicator 4.1.1 will not provide meaningful information about the relative quality of education systems, and the global distribution of educational opportunity.

Flexibility and rigidity: two problematic solutions

We could, of course, agree to disagree about the undefined terms in Indicator 4.1.1. In other words, each education system could interpret the terms in whatever way suits it, and report against the Indicator 4.1.1 using whatever measurement tools it has on hand. Alternatively, we could reach agreement by defining the terms with reference to a single test. In other words, all education systems that wish to report against Indicator 4.1.1 would be required to administer the same test. One of these solutions is extremely flexible, the other is extremely rigid; but neither of them is satisfactory.

The flexible solution has face-value appeal, since it would respect national authority and autonomy, as well as help to keep the costs of reporting against the SDGs to a minimum. But the flexible solution will not enable Indicator 4.1.1 to provide a framework for understanding global progress towards the SDGs, and would be unlikely to generate a constructive international discussion about what quality learning means. If reporting against Indicator 4.1.1 is entirely inconsistent, then we can only monitor progress of individual education systems in isolation, rather than progress on a global scale.

The rigid solution of administering a single test in all countries may appeal to those who value the rigour of consistent, psychometrically valid assessment. The developed world is already some distance along this pathway, with the widespread adoption of international tests such as the Programme for International Student Assessment (PISA), Trends in International Mathematics and Science Study (TIMSS) and the Progress in International Reading Literacy Study (PIRLS). Many developing countries also participate in cross-national assessments, at either international or regional level. These assessments may help address the definitional challenges that confront Indicator 4.1.1, as well as providing a wealth of accessible data and some level of quality assurance.

Yet while these cross-national tests are expanding their global reach, we cannot expect that participation in any one assessment program will be the right solution for all education systems. Most international tests have been developed in western economies, and may contain content or levels of difficulty that are inappropriate outside this cultural milieu. Education systems joining existing assessment programs must accept their definitions of reading and mathematics, without having had the opportunity to engage in the construction of those definitions. The auxiliary information provided by these tests is also designed with developed world policy issues in mind. To ensure that tests are relevant to their particular contexts, education systems need to decide for themselves what they test and how they test it. Requiring education systems to join an assessment program to report against Indicator 4.1.1 risks reducing the SDGs to a product that systems must buy into, regardless of its relevance, rather than a process of meaningful change.

Learning progressions: an innovative middle ground

The development of global learning progressions offers an innovative middle ground between these two undesirable extremes. A learning progression is a scale that defines the constructs that constitute educational progress in a particular domain (say, reading or mathematics). Learning progressions are directional, in that lower points on the scale represent less learning, and higher points represent more. Locations along the scale may be described numerically, as proficiency scores, or substantively, as proficiency descriptions. The proficiency descriptions make it clear what learners are expected to know and be able to do at designated levels on the scale, while the proficiency scores enable learning to be quantified against the scale.

In recent decades, the outcomes of tests have been increasingly reported against described scales (which have sometimes been referred to as learning progressions). However, to fulfil the need for global understandings of learning progress, it is necessary to create a learning progression that describes a construct independently of any particular assessment tool used to measure it. The difference between a test and such a learning progression may be likened to the difference between a ruler (a measurement tool) and length (a construct). Although different kinds of rulers may be used to measure length, these measurements are consistent because of the common understanding of length that informs their design.

Reaching agreement on a learning progression

Reaching agreement on common learning progressions in reading and mathematics will not be easy. It will require extensive consultation with members of the international education community, including leaders in cross-national assessments, learning domain and curriculum experts, and national curriculum, assessment and education policy teams from the widest possible range of countries. It will also require rigorous empirical work, drawing on existing curricula and assessment programs to map constructs and calibrate items to define the common scale.

Representatives from all countries need to be part of the conversation about measurement against Indicator 4.1.1. Leaders in the international assessment community must engage with education systems that may not have the means to attend global assessment forums, but whose stake in the SDGs is arguably the highest. This will help to ensure that the learning progressions are developed in a way that respects national sovereignty and cultural values, and that they generate a sense of ownership among all stakeholders who will benefit from their use.

The program of empirical work required to develop global learning progressions must also

involve the most diverse possible range of education systems and assessment programs. This includes programs conducted in a variety of languages, using a variety of assessment methods, and for learners at various levels of education. Drawing on data from a wide range of programs will strengthen the robustness of the scale development, and help to identify the differences and disconnections across assessment programs that need to be addressed.

It would be naïve to expect that perfect global agreement on the learning progressions can be reached. Yet history provides examples where sufficient agreement has been achieved to operationalise proficiency scales across diverse country contexts for specific assessment programs – such as in PISA, TIMSS and PIRLS, or more recently in the Pacific Islands Literacy and Numeracy Assessment, which spans countries at different stages of economic development and indeed reporting of outcomes across different stages of schooling on a single scale. This suggests that such consensus may also be possible for a common learning progression, which is not linked to any particular test. Where there is goodwill, and mutual interest in arriving at a fit-for-purpose result, shared understandings are possible. The conversation required to reach these understandings is itself beneficial, in providing an opportunity for constructive discussion about what improvement in learning may look like worldwide.

Balancing flexibility and rigour

The learning progressions maybe viewed with some suspicion by those who are eager to protect the greatest possible flexibility in reporting against the SDGs. It is true that the development of learning progressions will call all education systems to account for student learning against commonly defined learning outcomes, and enable differences in their progress to be more clearly shown. It is therefore essential that the learning progressions are accompanied by an active program of support for education system improvement, and that results reported against the learning progressions are contextualised within each system's challenges and opportunities.

The learning progressions may also be viewed with scepticism by those eager to uphold the highest possible standards of psychometric rigour in global education monitoring. It is true that common learning progressions will not provide the same comparability as a universal test – but that is not the point. Cross-national assessment programs will continue to be a valuable tool for education systems with the capacity and desire to invest in high levels of quality and consistency. The goal in developing the learning progressions is to provide a complementary tool for reporting against Indicator 4.1.1 that responds pragmatically to the diversity in priorities and practices that continues to exist in education systems around the world.

Developing the learning progressions will include the design of an appropriate quality assurance process for their use in reporting against Indicator 4.1.1. This will aim to ensure that data are collected in a timely, resource-efficient fashion, and are sufficiently valid to be fit-for-purpose. Quality assurance processes must also be developed collaboratively, to generate productive global dialogue about reasonable international expectations for the quality of learning assessments, alongside debates about the quality of education.

Beyond the SDGs

The usefulness of global learning progressions is not confined to reporting against Indicator 4.1.1. The learning progressions will encapsulate valuable, empirically-based information about how learning progresses in the relevant domains, which education systems may use to improve the quality of their curricula, teaching and learning, school resources, and assessment programs. Cross-

national assessments may also be strengthened through alignment with the progressions, and their connection to national curricula may be improved through the deeper understanding of learning progress that they provide.

The greatest benefit of the learning progressions may lie in shifting attention from tests to constructs – that is, from student test scores, to what students know and can do. The understandings of learning progress applied in assessment do not exist in a vacuum, but reflect the broader intentions of the education system about what students should have the opportunity to learn. This shift in thinking has many potential benefits beyond improving assessment, and may also assist policy-makers in making better use of test data to inform system improvement, and teachers in understanding what test results mean for curriculum and pedagogical practice. Indicator 4.1.1 thus becomes almost an afterthought, as the conversation shifts from comparison of assessment data, to what it means for teaching, learning and system improvement.

A risk worth taking

There are many risks involved in developing the learning progressions. The success of the endeavour depends upon the willingness of the world education community to accept common working definitions of reading and mathematics, and to locate themselves on a common scale. Political concerns will undoubtedly arise as the work progresses, as well as legitimate concerns about quality and validity as the scales are refined and applied. The challenge to meet resource demands is continually present, especially as Indicator 4.1.1 is only one of many measures of learning in the SDGs, all of which require fit-for-purpose strategies. Governments must move beyond the current tendency to cease investment at the point of measurement, and recognise the importance of investing in solutions that improve assessment data's meaningfulness and impact. These are nevertheless risks worth taking. The progress already made on developing learning progressions shows that international alignment of learning outcomes is not out of reach – nor is there a need to impose a universal measurement process to achieve a workable level of consistency. The SDGs have taken an important step in progressing the global conversation about education, and the questions that they raise about the meaning of quality are worthy of significant international attention. Learning progressions are a valuable tool in finding answers to these questions, and equipping the world to find new ways forward in the use of data to drive global educational improvement.