ASSESSING LEARNING PROGRESSION IN AN INTERNATIONAL LARGE-SCALE CONTEXT

T-1

Eckhard Klieme

Leibniz Institute for Research and Information in Education

Keywords: learning progression; international student assessment; mathematical knowledge

One of the major goals of educational assessment is to diagnose knowledge states and learning progression on a very detailed grain size, like across a teaching unit, e.g., on the Pythagorean Theorem or the concept of floating and sinking. This kind of work typically uses open-ended tasks, student answers to teacher questions, or documents of student work which can be coded for different kinds or stages of understanding, and scaled using partial-credit IRT.

Large-scale assessment, on the other hand, mostly uses multiple choice items, because they are easier to develop, administer, and scale in a multi-site or even cross-cultural context.

This paper reports on the assessment of student understanding in pre-, post-, and follow-up tests on quadratic equations, administered in a large international study of mathematics teaching and learning in 652 classrooms.

Tests were developed based on a framework specifying knowledge, skills, and understanding with respect to (a) prerequisites and (b) competent mathematical practice in the area of quadratic equations. Pre- and post-tests had no overlap. International pre-/post-tests were scaled using two-parameter IRT, item parameters were allowed to vary in case of large country DIF, and proficiency levels were established according to the standards of international studies (PISA), using item demand features identified by math experts.

Items were piloted, and 55 (30 pre-, 25 post-) were kept for the main study involving 15,323 students from 7 countries in Asia, Europe, and Latin America. For practical reasons, only multiple choice items were used. One country administered a follow-up test including some items from the pre- and post-test, so item-level change could be identified as well.

Four levels of proficiency were identified in the pre-test, five levels in the post-test. The level at the start very much determined the level at the end, echoing the strong pre-post-correlation. However, proficiency levels help understand what this means in terms of math knowledge: students who started on the pre-algebra level had a 2/3 chance of not even meeting the lowest proficiency level in the post-test. On the other hand, 94% of the students who reach at least a basic conceptual understanding of quadratic equations at the end of the unit started from the highest proficiency level in the pre-test.

From an assessment point of view, this paper offers a discussion on whether and how large-scale data can be used for diagnosing qualitative patterns in growth of student knowledge. From a subject-matter perspective, the paper provides detailed analysis of student knowledge in algebra, and explains why it is so hard to foster mathematical understanding for low-performing students. From a cross-cultural perspective, establishing measures of growth is a step forward in international assessment.