“MODELING AND MEASURING COMPETENCIES”
CONFERENCE, BERLIN, 24–25 FEBRUARY, 2011

In February 2011, I visited Berlin in order to attend a special conference on “Modeling and Measuring Competencies in Higher Education”, organized by the Humboldt University of Berlin and the Johannes Gutenberg University of Mainz. It was a special conference as it addressed this topic only over a period of two days and featured mainly invited (keynote) speakers. While the audience consisted mainly of German participants, leading researchers from universities and testing institutes in Germany, the U.S.A. and Australia presented their views and work. After three elaborate keynotes in the morning, the afternoon on both days consisted of panel discussions or a so-called “town hall meeting” in which three to five researchers introduced their work in 15 minutes, followed by a lively and interactive discussion. During lunchtime on Friday, there was a poster round in which 14 posters were presented which were mainly from Germany, with some from Finland. It was a very interesting and inspiring but also confronting experience. I would like to discuss two crucial controversies that were illuminated during this two-day meeting (the first day focused mainly on large-scale, high-stakes assessments while on the second day, we paid much more attention to the individual student in the assessment process):

1. Are we measuring the cognitive aspects of competencies or something more?
2. Curriculum validity versus professional/labor market validity.

These discussions, and the positions and frame of reference which researchers adopt in this respect (either large-scale (high-stakes) or focusing on the individual student), have, in my opinion, a large impact on both the modeling and measurement of competencies (or competencies), as well as on the impact such an assessment can or should have on the teaching and learning process.

THE FIRST DAY: LARGE-SCALE (HIGH-STAKES) MEASUREMENTS AT THE INSTITUTIONAL LEVEL

The openings words of Prof. Sigrid Blömeke informed the audience about a large German Ministry of Education and Research funding initiative called “Modeling and Measurement of Competencies in Higher Education”. This initiative stimulates new, creative but also fundamental research, emphasizing more evidence-based innovation in teaching and learning in (higher) education. Ninety-four proposals were submitted from various disciplines, showing the interest in and relevance of this topic and the necessity of carrying out research in this field.

© 2013 Sense Publishers. All rights reserved.
After the opening speech, the first day was filled with contributions from German and North American researchers, which the exception of the keynote speaker, Karine Tremblay, the Senior Survey Manager of the “Assessment for Higher Education Learning Outcomes (AHELO)” international project of the Organisation for Economic Co-operation and Development (OECD). Overall, this day was characterized by the view of assessment as a large-scale, high-stakes undertaking, for the purpose of comparing (or even ranking) institutions at a national or international level. Probably because of this purpose, all assessments discussed this day were written tests, often containing multiple-choice formats, for gauging the cognitive aspects of competence. This issue was heavily discussed in the panel meeting in the afternoon during a discussion of four German projects. The majority of these research projects defined competence by its “narrow definition”, described by Klieme and Leutner (2006) as “context specific cognitive dispositions that are acquired and needed to successfully cope with certain situations or tasks in a specific domain”. Blömeke adds to these cognitive aspects the importance of taking emotional/motivational aspects into account and therefore assesses competence by addressing not only students’ cognition, but also their beliefs. The decision to deal with cognitive aspects only was defended by arguing that these elements can be measured objectively through written items, which is almost inevitable in large-scale assessments used to compare institutions. However, an additional argument was made for “curriculum validity”: measuring those elements that are also present in the curriculum. The assessments were drawn up after consultations with faculty members from various higher education (HE) institutions about the content of their curricula. The involvement of the labor market was not an issue, as the labor market for HE graduates was argued to be too vague and too broad. The question which came immediately to mind was obvious:

If the purpose is to innovate higher education (see the funding initiative), then assessing only those elements that are present in the current curriculum will not stimulate innovation in the curriculum, will it?

Prof. Richard Shavelson, a well-known American assessment specialist, offers us a way out of assessing cognitive aspects only in large-scale tests: he provides technical insight into assessing competencies through written test items that are also authentic performance assessment tasks and representative of competencies used in the real world. He argues that the starting point for developing such an assessment should be a careful description of the criterion behavioral domain in real life (see also Shavelson, 2010), and thus not the current curricula of HE institutions. His technical presentation was enforced by Roger Benjamin on the second day of the conference, who discussed the Collegiate Learning Assessment (CLA), which is used all over the U.S. in HE institutions (see also Benjamin, Chun, & Shavelson, 2009). This is a large-scale, high-stakes assessment of critical analysis and evaluation, problem-solving, persuasive writing and the mechanics of writing through an array of written (but relevant to real life) performance tasks scored by trained computer systems. The degree to which written tasks can be called performance tasks can be debated, but Benjamin convinced me that these