SIGRID BLÖMEKE, OLGA ZLATKIN-TROITSCHANSKAIA,
CHRISTIANE KUHN AND JUDITH FEGE

MODELING AND MEASURING COMPETENCIES IN HIGHER EDUCATION: TASKS AND CHALLENGES

INTRODUCTION

Measuring competencies acquired in higher education has to be regarded as a widely neglected research field. The progress made in empirical research on the school system since the 1990s – for example, through large-scale assessments such as the Trends in International Mathematics and Science Study (TIMSS) and the Programme for International Student Assessment (PISA) and through a massive expansion of instructional research in general – has revealed that nothing comparable exists at the higher education level. This deficit can be traced back to the complexity of higher education and academic competencies. Not only is there a variety of institutions, programs, occupational fields and job requirements, but also the outcome is hard to define and even harder to measure. Thus, the existing research deficit is caused in part by the complexity that characterizes the academic competencies of undergraduate, graduate and doctoral students owing to the inter- and intra-national diversity of study models, education structures, teaching performances, etc.

In the context of a differentiated tertiary education system, assessing the development of competencies among students presents a methodological challenge. From this perspective, modeling and measuring academic competencies as well as their preconditions and effects set high thresholds. Another challenge is the question of a suitable criterion (e.g., future job requirements) that will help to evaluate the acquisition of competence. The requirements of possible job areas and of academics change constantly.

POTENTIAL

To review and structure the multi- and interdisciplinary field of higher education research, a comprehensive analysis of the international state of research on modeling and measuring competencies in higher education was conducted (Kuhn & Zlatkin-Troitschanskaia, 2011). The report is based on a broad documentary analysis in the form of a literature review and analyses of data (including secondary analyses), among others, in the form of a systematic keyword- and category-based analysis of the core research databases and publications. In addition, seven interviews were conducted with international experts on relevant topics. These enabled the authors to identify global tendencies and areas of

S. Blömeke, O. Zlatkin-Troitschanskaia, C. Kuhn & J. Fege (Eds.), Modeling and Measuring Competencies in Higher Education: Tasks and Challenges, 1–10.

innovative research in higher education. Overall, the report reveals research deficiencies in the modeling and measuring of competencies of students and graduates, especially in Europe.

At the same time, however, the report revealed that sustainable approaches to empirical higher education research exist (cf. the OECD feasibility study "Assessment of Higher Education Learning Outcomes," AHELO, or the studies in the context of TEDS-M, cf. Blömeke, Suhl, Kaiser & Döhrmann, 2012; Blömeke & Kaiser, 2012; Blömeke, Suhl & Kaiser, 2011). The "Teacher Education and Development Study: Learning to Teach Mathematics" (TEDS-M), carried out in 2008 under the supervision of the International Association for the Evaluation of Educational Achievement (IEA), was the first effort to measure higher education outcomes on a large scale using nationally- and internationally-representative samples (for more details see Blömeke in this volume). The challenges which had to be met with respect to sampling, response rates, reliability and validity, scaling and reporting at some points seemed unsolvable. Research perspectives had to be adjusted across academic disciplines, borders and locations.

The remarkable results of TEDS-M provided substantive indications of how to meet the challenges of higher education research. We learned for the first time on a large scale and from test data about the interplay of teaching and learning at universities, the interplay of various facets of professional competencies, about culture – or better philosophies of schooling – driving the development of the teacher education curriculum, the mediating influence of university educators, and so on (see, e.g., Blömeke et al., 2012; Blömeke & Kaiser, 2012; Blömeke et al., 2011). In addition, the study provided the first concept of benchmarks: what could be possible in higher education if specific circumstances, for example, in terms of entry selection, opportunities to learn or quality control mechanisms, were set in place. Such evidence did not exist prior to the study.

AN INTERNATIONAL CONFERENCE IN BERLIN – EXCHANGE AND INSPIRATION

Much research has to be done to reveal the structure of academic competencies and to make them accessible to assessment. A comprehensive understanding of higher education should include the assessment of domain-specific competencies as well as of generic academic competencies. With respect to the development and generalization of meaningful theories, it is important to focus on individual universities and their programs, and to include research on sometimes idiosyncratic features. The lesson learned from prior attempts in higher education research is that there is a need to create research communities among universities and disciplines and to take advantage of expertise gained in other countries.

The conference "Modeling and measurement of competencies in higher education" (www.competence-in-higher-education.com) hosted by the Humboldt University of Berlin and the Johannes Gutenberg University Mainz, provided an opportunity to do just that. The state of the research in this field was summarized from an international perspective and across academic disciplines. Speakers and

participants took part in an interdisciplinary discourse on various theoretical and methodological approaches to modeling competencies acquired in higher education and also reflected on the strengths and weaknesses of these approaches. They offered insight into the most important research projects currently being conducted and they identified state-of-the-art developments as well as future tasks.

Several controversies and challenges became apparent during the conference. Whereas most of the participants agreed on a definition of competencies as context-specific dispositions which are acquired and which are needed to cope successfully with domain-specific situations and tasks, there was an issue about the range of these dispositions. Should the term "competencies" include cognitive facets only or is it important to include attitudes as well? Insufficient response rates and panel mortality were mentioned as the main challenges, but the limitations of paper-and-pencil approaches to the complex issues surrounding the measurement of higher education outcomes were also of concern. Furthermore, only those competencies which can be measured with regard to psychometric criteria typically are regarded as relevant. Would this limit developments in higher education?

All in all, the conference served as an excellent platform for the exchange of research experiences and perspectives and, thus, provided incentive for a new funding initiative (see below). The conference results documented in this volume may instigate improvements in the higher education system. Such improvements can be implemented on the macro level, the institutional level and on the level of individual teaching processes.

A NEW FUNDING INITIATIVE – REASON AND GOALS

To close the research gap and encourage higher education in Germany to become internationally competitive, the funding initiative "Modeling and Measuring Competencies in Higher Education" (KoKoHs) was launched by the German Federal Ministry of Education and Research (BMBF) at the end of 2010. Apart from the development of competence models, KoKoHs focuses on generating appropriate measurement models and instruments. The funding initiative is intended to provide incentives for basic competence research in the tertiary education sector. It has the following goals:

- To increase the performance of the German tertiary education system
- To keep up with international competence research in higher education
- To develop a foundation for the evaluation of competence development in higher education so that evidence-based policy decisions can be made.

In particular, the initiative is intended to support innovative research projects striving for cooperation among universities. The announcement of the funding initiative elicited 97 high-quality proposals for modeling and measuring competencies: in engineering; economics; education and psychology; teacher education in science, technology, engineering and mathematics (the STEM subjects); and social sciences, as well as generic academic competencies. These fields were selected as priority areas where research needs to start for synergetic

effects to be optimized. After an evaluation conducted according to the criteria of the German Research Foundation (DFG), about 20 research projects were selected. They will receive funding from the end of 2011 or beginning of 2012 until the end of 2014. Experts from various disciplines will work together and network nationally as well as internationally in joint multi- and interdisciplinary research projects while integrating diverse methods. The projects are expected to pay attention to certain – almost quasi natural – areas of conflict, for example, the tension between curricular validity, job requirements and the dynamics of changing labor markets in a globalized world.

Proactive funding initiatives based on deficit analyses and aimed at developing a new field of research often face the problem – if one insists on funding according to quality assessments – that only a small number of submissions can be reviewed positively (cf. Nießen, 2011). In the context of earlier initiatives, the federal ministry noticed to its chagrin that financial constraints were not the limiting factor in terms of funding: the quality of proposals was simply not high enough. However, with the new funding initiative on higher education, the picture has started to change and even high-quality applications had to be rejected. This can be regarded as an important signal of the increasing competitiveness of higher education research.

Coordination offices were opened on May 1, 2011 in Berlin (under the direction of Sigrid Blömeke, Humboldt University of Berlin) and Mainz (under the direction of Olga Zlatkin-Troitschanskaia, Johannes Gutenberg University Mainz) to administer the projects and the research program. The coordination offices strive to create a systematic framework for the individual projects and a structured approach, aiming to reach the ultimate goals of the program by developing a superordinate concept. The main tasks of the coordination offices are to cultivate exchange and networking opportunities among the projects being promoted, to use synergies, and to foster the systematic and sustainable promotion of young scientists. A special concern is to maintain international cooperation and use it for exchanging communication within the national funding initiative. The coordination offices are expected to remain open for four years so KoKoHs can be supervised during the complete funding period.

OVERVIEW: THE PAPERS IN THIS VOLUME

The conference and the funding initiative will contribute significantly to the advancement of higher education research. Few other factors are as important to sustainable human progress, social justice and economic prosperity as the quality of education – and it is the responsibility of researchers to contribute by conducting high-quality studies, the results of which will lead to improved understanding of the processes and outcomes of teaching and learning. Laying the foundation for this outcome in the field of higher education was the core aim of the conference. Each talk and poster focused on a pressing issue in this field and the conference – with 230 prestigious participants – was an excellent two-day learning experience and, thus, the conference achieved its aim. The interdisciplinary pool of 16 speakers from the Americas, Australia and Europe reached the conclusion that

there are theoretical and methodological approaches to modeling and measuring competencies in higher education that are worth pursuing.

Part I: Theory and Methodology

Royce Sadler, a Professor at Griffith University in Brisbane, Australia, specializes in formative assessment theory and practice, discusses the term "competence" in his paper "Making competent judgments of competence." He points out that the term "competence" differs only slightly in spelling from "competency" but that there is a conceptual distinction between them which in turn leads to distinct approaches to their measurement. A "competency" is often taken to mean an identifiable skill or practice. "Competence," in contrast, is often understood to consist of a large number of discrete competencies which can be tested independently by objective means. Competence involves being able to select from and then orchestrate a set of competencies to achieve a particular end within a particular context. The competent person makes multi-criteria judgments that consistently are appropriate and situation-sensitive. What is more, the range of situations faced by many professional practitioners is potentially infinite. Dividing competence into manageable components to facilitate judgment has value in certain contexts, but the act of division can obscure how a practitioner would connect the various pieces to form a coherent whole. Sadler makes a plea for more integrative and holistic judgments to arrive at consistent evaluations.

Richard Shavelson, Professor (Emeritus) at Stanford University in the US and a specialist on the measurement of human performance, presents an interesting approach to testing and modeling competency. He describes competency as a complex ability closely related to real-life-situation performance. How to make it amenable to measurement is exemplified by research from the business, military and education sectors. Generalizability, a statistical theory for modeling and evaluating the dependability of competency scores, is applied to several of these examples. In his paper he then puts the pieces together in a general competency measurement model. Shavelson points out, however, that there are limitations to measuring competency on various levels in terms of resources, costs and time.

Fritz Oser, Professor (Emeritus) at Fribourg University in Switzerland and a specialist in developing standards for teacher education, in his paper "Competence Profiles" emphasizes the process of generating criteria against which competence can be evaluated. He claims that basic questions on professionalization background and the identification of standards have to be answered before competence profiles at the university level can be modeled and assessed. Oser demonstrates how the Delphi method can identify vital competencies. He has developed an advocatory approach to measuring competencies based on the assumption that the individual situation defines the competence profiles which, therefore, should be defined from the bottom up. He presents corresponding results from his study.

Mark Wilson, Professor at the University of California, Berkeley (USA), and Karen Draney, specialists in educational measurement and psychometrics, focus on an assessment system which has been developed by the Berkeley Evaluation and

Assessment Research (BEAR) Center. They briefly describe a large-scale assessment context in which they have been developing and applying aspects of the BEAR Assessment System. They describe BEAR in terms of its principles and building blocks and discuss its realization in their large-scale context. Throughout their paper they discuss what their experiences have taught them regarding some of the salient issues regarding assessment.

Michaela Pfadenhauer, Professor at the Karlsruhe Institute of Technology in Germany and a specialist in the sociology of knowledge, in her paper "Competence – more than a buzz phrase and an emotive word?" examines the evolving use of the term *competence* as an indicator of changing educational systems. She points out that in educational policy – at both the national and the supranational level – a "competency-oriented turn" has taken place on such a scale that it is hardly conceivable how it was possible to manage without this phrase. Its rise in popularity was accompanied by a massive replacement of customary concepts: where "qualification," "education" and "educational objectives" previously were discussed, "competency" now seems to be the more accurate, adequate or simply more modern expression. Pfadenhauer takes a perspective on situational problem-solving capacity; on the basis of her phenomenological analysis, she makes a plea for including the social dimension in the definition of competencies.

Part II: Instruments and Studies

Sigrid Blömeke, Professor at the Humboldt University of Berlin, a specialist in the measurement of teacher competence and one of the conference organizers, presents an innovative comparative study carried out under the supervision of the International Association for the Evaluation of Educational Achievement (IEA): the "Teacher Education and Development Study: Learning to Teach Mathematics" (TEDS-M). In her paper she describes the theoretical framework of this large-scale assessment and its design to illustrate how the challenges of higher education research were met. Core results of TEDS-M are documented to illustrate the potential of such studies. Finally, conclusions are drawn with respect to further higher education research.

Karine Tremblay, Senior Survey Manager, Organisation for Economic Co-operation and Development (OECD), France, a specialist in statistics in the areas of student mobility and assessment of learning outcomes in higher education, presents the rationales, challenges and insights derived from OECD's feasibility study "Assessment for Higher Education Learning Outcomes" (AHELO). AHELO is intended to provide evidence of outcomes across cultures and institutions for national and international use in developing policies and practices in higher education. AHELO targets discipline-related competencies and generic skills (critical thinking, analytic reasoning, problem-solving, written communication). In contrast to other OECD studies such as PISA, the unit of analysis is not the country but the institution. Feedback is obtained through performance profiles. Major research questions of the feasibility study are whether instruments are valid in diverse national and institutional contexts, whether the tests meet predefined

psychometric standards and how effective strategies are in encouraging institutions and students to participate.

Roger Benjamin, President of the Council for Aid to Education (CAE) in the USA and a specialist in higher education policy and practice, examines "the principles and logic of competency tests for formative and summative assessment in higher education." He starts his paper with a reminder of the reason why such efforts are made: the future of our highly-industrialized society depends on the realization of human capital. Therefore, a need exists for evidence-based decisions focused, in particular, on the improvement of teaching and learning. Benjamin presents the "Collegiate Learning Assessment" (CLA) as an approach to capturing one key competence to be developed in higher education: critical thinking. In his paper, he presents the lessons learned in developing and adapting this performance assessment instrument for international use. The CLA requires students to use their cognitive abilities to construct responses to realistic problems. Benjamin also addresses an important concern: that taking a test has to be more enjoyable than going to the dentist.

Rafael Vidal Uribe, Director of the National Assessment Center for Higher Education (Ceneval) in Mexico and a specialist in large-scale assessments, presents "The case of Ceneval in Mexico" as an example of measuring learning outcomes in higher education. Two main instruments are used to evaluate college graduates. The EXANI-III evaluates the fundamental skills and competencies of those who have completed college and wish to continue with post-graduate studies. The EGEL examinations are designed to assess the required knowledge expected of scholars on completion of their first degree studies. The EGEL examinations are multiple-choice tests centered on the domain-specific knowledge and skills that are considered essential and common to all higher education institutions' curricula in the specific subject. The objective is to identify whether students have the minimum knowledge, skills and competencies they need to enter professional practice. Results for individual students are reported on one of three levels (outstanding, satisfactory, not yet satisfactory) and described on each subscale. Results for the institutions are reported through the distribution of students on the three levels for each subscale across all subjects.

Hildegard Schaeper, Senior Researcher at the Institute for Research on Higher Education (HIS), is involved in Stage 7 (Higher Education and the Transition to Work) of the German National Educational Panel Study (NEPS) and is responsible for project coordination and management. In her article, she first gives a brief overview of the conception and structure of the NEPS and then describes in more detail its general approach to modeling and measuring competencies and its method of addressing the issue of subject-specific competencies in higher education. The NEPS promises to gain new insights into the acquisition of competencies across the entire lifespan, to describe crucial educational transitions, to study educational careers, to identify the determinants of competence development and educational decisions, and to analyze the impact of education and competencies over the life course.

Olga Zlatkin-Troitschanskaia, Professor at Johannes Gutenberg University Mainz, and Manuel Förster and Christiane Kuhn specialize in the measurement of university students´ competence in the domain of business and economics. As one of the conference organizers, Zlatkin-Troitschanskaia presents the research project ILLEV. It is one of the few projects in the German Federal Ministry of Education and Research's funding program "University Research as a Contribution to Professionalizing Higher Education" that focuses on modeling and measuring subject- and subject-didactical competence, especially among students of business and economics and business and economics education. In the study, the effects of the various courses of study (diploma and bachelor/master) on professionalization and its development over the course of four years are examined. After discussing the study's basic aims and research questions, the research design, and the survey instruments employed, this paper provides a description of the main content and measuring results of the first survey (fall 2008). The paper concludes with a discussion and preview of further approaches in this longitudinal study.

Detlev Leutner, Professor at the Duisburg-Essen University in Germany, and Karoline Koeppen, Johannes Hartig and Eckhard Klieme present the program "Competence Models for Assessment of Individual Learning Outcomes and the Evaluation of Educational Processes." This priority program, which is based on proposals written by individual researchers, was set up by the German Research Foundation (DFG) to operate for six years (2007–2013). It coordinates the research of experts on teaching and learning as well as experts on measurement and assessment from the disciplines of psychology, educational science and domain-specific pedagogics in more than 20 projects across Germany.

Part III: Long-term Outcomes

Christiane Spiel, Professor at the University of Vienna in Austria, together with Barbara Schober and Ralph Reimann, specialists in evaluation and quality management in the educational system, stresses the institutional perspective. She focuses on "The Contribution of Scientific Evaluation" to the measurement of academic competencies. Scientific evaluation is based on established standards and systematically combines qualitative and quantitative approaches to data collection and analysis. Spiel makes the plea that evaluation can and should be conducted in all phases of programs and from a longitudinal perspective. Baseline data collected before the start of a program are used to describe the current situation, for example, the generic and domain-specific competencies of students before beginning their university education. In formative evaluation, interim data are collected after the start of a program but before its conclusion. It is the purpose of formative evaluation to describe the progress of the program and, if necessary, to modify and optimize its design. In the case of higher education, the focus might be on how academic study and specific courses support the development of generic and domain-specific competences. Outcome evaluation deals with the question of whether programs achieve their goals. Here, the generic and domain-specific

competences of graduates and freshmen (baseline data) can be compared. Furthermore, the competences of graduates might be evaluated in relation to their correspondence to defined profiles.

Rolf Van der Velden, Professor at Maastricht University in the Netherlands and a specialist in the long-term effects of education on careers, stresses the ultimate criterion of competence acquired during higher education leading to success in life, especially in the labor market. He makes a plea for including non-cognitive facets in this evaluation. Drawing on this background, he discusses two of the main methods of measuring competencies in large-scale surveys among higher education students or graduates: tests, and self-assessments.

Ulrich Teichler, Professor (Emeritus) at the University of Kassel in Germany, and Harald Schomburg, both specialists in the internationalization of higher education, analyze job requirements and the competencies of graduates. Teichler points out that even though the measurement of competencies can be regarded as the most sophisticated approach to evaluating the quality of higher education, drawbacks may exist. Higher education research has to identify the key actors' notions of job requirements and competencies of graduates, that is, the notions of employers, students and academics. He introduces the term "subversity," albeit as a safeguard against the mostly conventional ideas of employers and university professors. Four areas are most salient if improvement is to be achieved: (a) concepts are needed to overcome the "match-mismatch" paradigm, that is, to take into account the necessary concurrent "over-" and "under"-education, the educational tasks beyond professional preparation, the varied values of graduates, the creative function of presumed "over-education," etc.; (b) methods have to become better at de-mystifying misconceptions between job requirements and competencies; (c) ways have to be found to create a better balance between subject-related competencies (e.g., mathematical reasoning) and general competencies (e.g., leadership); and (d) it is still an open question how one should measure competencies and job requirements in such a way that the varied demands in the employment systems and the varied curricular concepts in higher education are taken into serious consideration.

Part IV: Commentary

Judith Gulikers and Martin Mulder took on the task of summarizing and commenting on what was to be learned at the conference from the participants' point of view. They relate the ideas presented in Berlin, among others, to research work done in the Netherlands and, thus, pave the way for an even broader view of measuring competencies in higher education. In particular, they identify the challenges ahead if we are serious about moving forward in this research field.

As the conference organizers and editors of this volume, we are grateful for the contributions of all our speakers and participants. Special thanks go to the members of our Advisory Board, in particular to its head, Prof. Dr Klaus Beck. The Board members supported us with great ideas and recommendations, and also actively participated in the conference by introducing the speakers and leading the

discussions. We are grateful for the support of Manuel Förster, Sebastian Brückner and Katharina S. Bergsma as well. They worked tirelessly prior to, during and after the conference so that it ran smoothly, guests felt welcome and this volume could be issued on time. All contributions were subject to double-blind reviews; therefore, we would like to thank all colleagues who contributed to the reviewing process. Finally, we gratefully acknowledge the funding provided by the BMBF represented by Martina Diegelmann, Michael Kindt and Hartung Hoffmann, who are also in charge of administering the funding initiative. The conference has revealed how complex the task of measuring academic competencies is and that there is a lot of research work to be done. We anticipate, however, that we will move forward substantially over the next three years and beyond – thanks to the more than 20 research projects in this initiative.

Berlin & Mainz, Germany, January 2013

REFERENCES

Blömeke, S., & Kaiser, G. (2012). *Homogeneity or heterogeneity*: Profiles of opportunities to learn in primary teacher education and their relationship to cultural context and outcomes. ZDM – The International Journal on Mathematics Education. DOI 10.1007/s11858-011-0378-6.

Blömeke, S., Suhl, U., & Kaiser, G. (2011). Teacher education effectiveness: Quality and equity of future primary teachers' mathematics and mathematics pedagogical content knowledge. *Journal of Teacher Education*, *62*(2), 154–171.

Blömeke, S., Suhl, U., Kaiser, G., & Döhrmann, M. (2012). Family background, entry selectivity and opportunities to learn: what matters in primary teacher education? An international comparison of 15 countries. *Teaching and Teacher Education* 28, 44–55.

Kuhn, C., & Zlatkin-Troitschanskaia, O. (2011). *Assessment of competencies among university students and graduates – Analyzing the state of research and perspectives*. Johannes Gutenberg University Mainz: Arbeitspapiere Wirtschaftspädagogik [Working Paper: Business Education], 59.

Nießen, M. (2011). Building structures by research funding? DFG programmes in the field of empirical research in education. Zeitschrift für Erziehungswissenschaft, *Sonderheft*, *13–2011*, 161–169.