

# Valid Competency Assessment in Higher Education: Framework, Results, and Further Perspectives of the German Research Program KoKoHs

**Olga Zlatkin-Troitschanskaia**

*Johannes Gutenberg University Mainz*

**Hans Anand Pant**

*Humboldt-Universitaet zu Berlin*

**Miriam Toepper**

*Johannes Gutenberg University Mainz*

**Corinna Lautenbach**

**Dimitar Molerov**

*Humboldt-Universitaet zu Berlin*

*The aim of the 15 collaborative projects conducted during the new funding phase of the German research program Modeling and Measuring Competencies in Higher Education—Validation and Methodological Innovations (KoKoHs) is to make a significant contribution to advancing the field of modeling and valid measurement of competencies acquired in higher education. The KoKoHs research teams assess generic competencies and domain-specific competencies in teacher education, social and economic sciences, and medicine based on findings from and using competency models and assessment instruments developed during the first KoKoHs funding phase. Further, they enhance, validate, and test measurement approaches for use in higher education in Germany. Results and findings are transferred at various levels to national and international research, higher education practice, and education policy.*

Keywords: *student learning outcomes, assessment, test instruments, validation, methodological innovations, higher education*

## Assessing Student Learning Outcomes in Higher Education in Germany: Background of the KoKoHs Program

SINCE reforms were implemented that introduced outcome orientation in higher education, such as the Bologna reform, policy discussions on the quality of teaching in higher education in Germany increasingly have focused on students' competencies as key learning outcomes (see Standing Conference of the Ministers of Education and Cultural Affairs of the Federal Republic of Germany; Kultusministerkonferenz, 2000). In Germany, outcome orientation focusing on students' competencies, understood as complex abilities and skills to be acquired in education, has implied a central claim: Higher education not only should provide students with the general opportunity to attain top educational achievement in any given discipline—this ideal of academic education dating back to Wilhelm von Humboldt still has a very strong appeal today—it should prepare them for challenging professional positions and

enhance their employability. Consequently, in the academic teaching-and-learning process, students are expected to acquire scientific competencies related to their field of studies and develop the ability to reflect on and expand their knowledge and apply and adapt it to various areas of professional practice. In addition, higher education is expected to promote the acquisition of transdisciplinary, multifunctional skills and key competencies. According to recommendations by the German Council of Science and Humanities (Wissenschaftsrat, 2015), higher education institutions have the responsibility to address and promote the three central goal dimensions of academic education—research and study, personality formation, and preparation for the job market. These aims and expectations are in line with current definitions and provisions in frameworks at the international level, such as the European Qualifications Framework (EQR) (see also European Commission, 2014; for a competency-based education model in the United States, see e.g., Soares, 2012).



In this light, competency orientation calls for new approaches to teaching in higher education that value learning outcomes over study content. Focus is shifting from content and knowledge to complex abilities and skills. With regard to higher education practice, the idea of competency orientation should be understood as a system of ongoing improvement of teaching and learning in which curriculum, instruction, and assessment need to be brought into optimal alignment, as described in the model of constructive alignment (see Biggs & Tang, 2011). Examinations and competency assessments represent a key element of this triad (see Pellegrino, Chudowsky, & Glaser, 2001). Hence, examination procedures with feedback systems are needed to provide teachers and learners with reliable information on whether and to what extent teaching-and-learning objectives are being met and how learning opportunities can be enhanced.

Competency orientation requires implementation of new measures at various levels and in various areas of practice. Measures could include, for instance, changing teaching-and-learning constellations or offering training courses in teaching methods for higher education teachers. Curricula need to be reconceptualized based on qualification objectives, competency standards need to be formulated, and examination formats need to be tailored accordingly.

These tasks are complex and become challenging in view of the heterogeneous and multidimensional requirement profiles for graduates and the diverse study models and programs offered by higher education institutions. Competency-oriented study objectives have been introduced formally; study programs now need to be developed further. For this purpose, theoretically sound models of students' generic and domain-specific competencies as well as methodologically adequate instruments for assessing these skills and abilities are needed. Such instruments would serve to assess individual performance and would be valuable tools for monitoring learning progress, implementing quality assurance measures to evaluate teaching, and redesigning learning opportunities based on evidence. Assessments provide knowledge for decision making for university management staff, policymakers, and administrators, for example, with regard to questions of credit approval and admissions, which are becoming more important with the internationalization of higher education, but also with regard to the effectiveness and efficiency of teaching in higher education in general (see also Coates, 2016).

To assess students' competencies in a transparent and objective way so that they can be compared, suitable examination and certification systems for higher education practice are needed. In view of current societal developments affecting higher education, such as massification of degree courses, internationalization of study programs, migration, and global mobility of students, the increasing diversity of student bodies must be taken into account when redesigning higher education to promote competency acquisition. The

share of students who enroll in higher education globally has been on the rise each year for the past 20 years (see OECD 2015), leading to a growing number of mass courses, more diversity in the classroom, and less time for teachers to address needs of individual students. Furthermore, the internationalization of study programs and global mobility of students call for greater transparency of and valid information on students' knowledge and skills. Several theoretical and methodological challenges arise from the immense diversity within student bodies, degree courses, study programs, and institutions (see Zlatkin-Troitschanskaia, Shavelson, & Kuhn, 2015).

These resulting challenges pertain to all study domains and are summarized in a key question: How can higher education institutions truly embrace diversity by proactively and effectively tapping the potentials of diversity and multinationalism while offering tailored study programs and learning opportunities for diverse groups of students? Suitable instruments are needed in order to heed the principle that good therapy requires good diagnosis. To this end, testing procedures suitable for diagnostic purposes need to be implemented in order to foster students' competencies, that is, their ability to perform in real-life and professional situations. Accordingly, test items should be developed that challenge students to apply their knowledge and skills in realistic, action-oriented contexts. A guiding principle in this regard is that a good test item also is a good exercise task. Schneider (2009) emphasizes that "as educators, we need to move beyond the reactive mode provoked by the Spellings barrage and help society get ahead of the curve on forms of assessment that can actually drive higher achievement" (p. 2). Designing assessments for process diagnostics of such competencies or complex skills is a conceptually and methodologically highly demanding task and requires considerable effort. This certainly is one of the reasons competency orientation is so difficult to implement in practice.

In the past decade, valid assessment of students' competencies and their development over the course of their studies has been central to many international research projects in which the conditions, design, and effects of teaching and learning in higher education are examined (see also Coates, 2014, 2016). However, a review of the state of international research and assessment practice in OECD countries indicated that there still are very few substantial and adequately elaborate research projects and studies in which valid assessment of competencies of higher education students are examined (see Zlatkin-Troitschanskaia et al., 2015). The lack of research and available assessments in educational practice has prompted researchers to pay more attention to the acquisition of competencies, influences on competency acquisition, and ways to promote competency acquisition and systematically enhance the quality of existing assessments. Valid assessment of competencies in higher education is the basis for establishing comparable academic

degrees, which is a stated policy objective of educational reform programs. To this end, in 2010, the German Federal Ministry of Education and Research (BMBF) initiated the national Modeling and Measuring Competencies in Higher Education (KoKoHs) research program, which addresses the political and practical challenges of conducting competency assessments in higher education.

### **Modeling and Measuring Competencies in Higher Education: Aims of the First Phase of the KoKoHs Research Program**

The KoKoHs program was funded as part of the BMBF funding area Research on Science and Higher Education. The general aim of the funding has been to build new and strengthen existing research and development capacities in the areas of science and higher education. Initial research findings on valid assessment of competencies in the higher education system in Germany recently have been transferred into practice, and the effects have just started to gain recognition among practitioners involved and educational policymakers. In view of the above developments, more research is needed to produce scientifically substantiated findings on the acquisition of competencies in higher education as well as their preconditions, effects, and measures for optimization. The output of the higher education system needs systematic improvement. After successful advances in research on competencies in secondary education (e.g., see Leutner, Fleischer, Grünkorn, & Klieme, in press), policymakers recently have started to contemplate how communication and decision-making processes can be informed by valid assessment of competencies acquired in higher education.

During the first funding phase of the KoKoHs program from 2010 to 2015, focus was on fundamental research and development of theory-driven models of generic and domain-specific competencies and of corresponding assessment instruments in various study domains. The development of instruments involved operationalizing theoretical models and identifying and describing competency and learning requirements from degree course profiles and typical areas of professional practice (see Zlatkin-Troitschanskaia et al., 2017). KoKoHs project teams took into account both curricular and job-related requirements, defined constructs to be assessed based on the international literature in the respective fields and in psychometrics, and specified levels, content, and cognitive requirements in competency models according to the general concept of competencies adhered to in the program. In KoKoHs, competencies were defined as latent cognitive and noncognitive underpinnings of performance (see Ewell, 2005; Rychen, 2004; Shavelson, 2013). Weinert's (2001) definition of competencies as being "cognitive abilities and skills that individuals possess or acquire in order to solve certain

problems as well as the aligned motivational, volitional and social dispositions and skills to apply the solutions in different situations successfully and responsibly" (pp. 27–28) was adopted for higher education contexts. During the first phase, KoKoHs project teams focused predominantly on (latent) cognitive abilities and skills and modeled them for various fields of study.

Subsequently, KoKoHs project teams developed model-based measuring instruments and validated their test score interpretations according to the Standards of Educational and Psychological Testing ("the Standards," American Educational Research Association [AERA], American Psychological Association [APA], & National Council on Measurement in Education [NCME], 2004, 2014) so as to enable users of the tests to draw evidence-based conclusions about outcomes of higher education in Germany (see also Kane, 2013; Tiffin-Richards & Pant, in press). In addition to systematically modeling and assessing generic and domain-specific competencies of higher education students and graduates, KoKoHs project teams examined numerous factors determining the level of students' competencies, including institutional variables related to issues of institutional performance and policy and personal variables such as gender, sociocultural background, and prior knowledge related to issues of equity, heterogeneity, and connections to previous and subsequent educational stages. With this broad focus, the KoKoHs program has built capacity and academic infrastructure on a national scale, and KoKoHs research teams have conducted internationally compatible fundamental research on competency assessment in higher education in Germany (Zlatkin-Troitschanskaia, Kuhn, & Toepper, 2014).

### **Program Structure and Projects**

During the first phase, the KoKoHs program comprised 24 collaborative, cross-university projects, encompassing 70 single projects with 220 researchers at more than 50 institutions of higher education in Germany and Austria. Each collaborative project brought together content domain experts, teaching methodology experts, and research methodology experts from at least two universities. The 24 collaborative KoKoHs projects were sorted into five clusters according to the field of study being assessed, including four clusters for domain-specific competencies and one cluster for generic competencies.

The selected fields of study included economics, engineering, educational sciences, and teacher training in science, technology, engineering, and mathematics (STEM subjects) (for examples of projects, see the WiWiKom project for economics, Modeling and Measuring Competencies in Business and Economics among Students and Graduates, Zlatkin-Troitschanskaia, Förster, Brückner, & Happ, 2014; the KOMING project for engineering, Modeling and Measurement of Competencies of Engineering Mechanics in the Training of

Mechanical Engineers, Musekamp, Spöttl, Mehrafza, Heine, & Heene, 2014; the KoWaDis project for teacher training in STEM, Evaluating the Development of Scientific Literacy in Teacher Education, Hartmann, Upmeier zu Belzen, Krüger, & Pant, 2015). Domain-independent, generic competencies acquired in higher education that were assessed in KoKoHs included, for example, students' competencies in dealing with academic texts or self-regulating during their studies (see, e.g., the LeScEd project for generic competencies, Learning the Science of Education, Groß Ophoff, Schladitz, Leuders, Leuders, & Wirtz, 2015). In the respective domains, international approaches to competency assessment were reviewed and if possible adapted and further developed to create suitable instruments for use in Germany (see Brückner, Zlatkin-Troitschanskaia, & Förster, 2014). If available, approaches from secondary education in the respective fields in Germany also were taken into account if they were domain-specific and had an empirical basis (see, e.g., Obersteiner, Moll, Reiss, & Pant, 2015).

KoKoHs project teams have been cooperating with researchers involved in other relevant educational assessment programs in Germany such as the National Educational Panel Study (NEPS). KoKoHs has promoted the establishment of a new research community in the area of competency modeling and measurement in higher education and an international network of leading institutions and researchers in this area. KoKoHs project teams have been cooperating with more than 50 international experts (from universities, testing institutes, etc.) from 20 countries, such as the United States, Australia, Japan, South Korea, and Mexico. During the first phase, international cooperation comprised joint research analyses, for example, in the WiWiKom project (see Zlatkin-Troitschanskaia, Förster, et al., 2014), organization of joint events (see, e.g., Kuhn, Toepper, & Zlatkin-Troitschanskaia, 2014), preparation of joint publications (see, various special issues, e.g., Pant & Zlatkin-Troitschanskaia, 2016; Zlatkin-Troitschanskaia & Shavelson, 2015), and promotion of young researchers (see, e.g., Toepper, Zlatkin-Troitschanskaia, Kuhn, Schmidt, & Brückner, 2014).

To achieve the stated aims of the KoKoHs research program and create a basis for reliable and valid assessment of students' generic and domain-specific competencies in higher education in Germany, the following three major milestones and areas of work had to be reached.

First, competency models were defined, that is, knowledge, skills, and abilities to be taught and acquired in higher education were described based on teaching methodology and learning psychology. Second, the models were operationalized into test instruments to assess the competencies acquired in different fields of study at various stages of studies in higher education, including during the entrance phase, over the course of studies, and during transition into the profession. Third, the assessments were tested and validated for use in

higher education in Germany according to the Standards (AERA et al., 2004; 2014). A detailed overview of the program design, assessment framework, developed competency models, test instruments, and validation results recently was published by Zlatkin-Troitschanskaia et al. (2017).

In the following, an overview of the main outcomes of the first KoKoHs funding phase is presented, which involved competency modeling, test development, and validation.

### **Outcomes of the First Phase of the KoKoHs Program: Advances and Frontiers**

The first milestone was reached during the first KoKoHs funding phase with the development of 40 competency models defining diverse generic and domain-specific competencies. To ensure content validity of the models, which referred mostly to curricular validity, KoKoHs project teams analyzed altogether approximately 1,000 documents on learning objectives (e.g., module descriptions and study regulations) from more than 250 higher education institutions throughout Germany as well as 1,500 documents on existing items (e.g., from exams, exercises, and lecture notes), project reports, and lab reports. In addition, interviews with more than 500 experts and cognitive interviews with almost 500 students were conducted as part of the validation process of the assessments developed (for a detailed overview, see Zlatkin-Troitschanskaia et al., 2017).

The second milestone was reached with the operationalization of the competency models into assessment instruments (for examples of tests, see Zlatkin-Troitschanskaia et al., 2017). Overall, more than 60 paper-pencil tests, almost 40 computer-based tests, and 10 video-based tests were developed in KoKoHs. For some competency facets, more action-oriented approaches were developed, such as videographed role plays or computer-based learning diaries (for a detailed overview, see Zlatkin-Troitschanskaia et al., 2017).

Table 1 provides an overview of the study phases for which the newly developed tests are suited (i.e., beginning, middle, and end of bachelor or master studies and practical phase in teacher training).

Table 2 indicates the geographical scope of instruments, that is, whether they have been used in local assessment, large-scale assessment, or international assessment; details are given on the number of items and countries where they have been used.

The third milestone was reached with the testing of the newly developed assessments in field studies with approximately 50,000 students from almost 230 higher education institutions throughout Germany. Table 3 shows the overall sample sizes for each cluster and the number of students tested as well as their focus of studies.

These initial field studies provided evidence of the quality of the competency models and instruments. In the validation studies, project teams adhered to the Standards (AERA

TABLE 1  
*Area of Application of KoKoHs Instruments According to Cluster*

Area of application	Generic competencies cluster	Teacher training STEM cluster	Educational science cluster	Economics and social sciences cluster	Engineering cluster
Beginning of studies	4	6	4	2	3
Course of studies	6	3	4	3	3
End of studies, bachelor	6	6	5	3	3
End of studies, master	6	5	4	3	3
Professional practice		1	2	2	

et al., 2004, 2014) and conducted comprehensive analyses of the internal structure of the assessed constructs and their relationship to other variables such as prior knowledge, grades, and so on (for examples, see Brückner, Förster, Zlatkin-Troitschanskaia, Happ, et al., 2015; Brückner, Förster, Zlatkin-Troitschanskaia, & Walstad, 2015; Groß Ophoff et al., 2015; Hartmann et al., 2015). Apart from instrument validation, the field studies also served to generate research findings on students' competency levels in the selected domains and on structural and individual influence factors (for more detailed results from the KoKoHs projects, see various special issues on KoKoHs findings, e.g., Zlatkin-Troitschanskaia & Shavelson, 2015).

While the main aim of the first KoKoHs funding phase was to provide empirically tested and validated competency models and instruments, results of these fundamental studies have been implemented in higher education practice in various ways. A large number of higher education institutions that participated in the assessments have used the results of the projects to design competency-oriented courses, tasks, tests, and examinations. New learning opportunities, tutoring, and training have been created to foster various facets of students' generic or domain-specific competencies (e.g., explaining skills of pre-service teachers); they have been evaluated using KoKoHs tests and have been included in the study curriculum. Some new learning opportunities have resulted directly from KoKoHs studies such as a course on case study-based teaching in laboratories for science subjects at school. Empirical evidence from KoKoHs projects is being used to revise curricula and learning goals with regard to the competencies to be acquired. Some KoKoHs instruments also have been used for evaluation and accreditation of degree courses, for example, in the educational sciences. A few examples from the cluster of teacher training in STEM subjects illustrate the transfer activities that have taken place and their impact on practice. A short test developed in KoKoHs for individual diagnostics has been used at a university to examine the development of STEM teaching competencies in a graduate program and has been requested by other higher education institutions. KoKoHs findings on the importance of practical learning in laboratories in teacher

training for STEM subjects have prompted discussion among experts in subject matter and teaching methodology departments at the institutions assessed, ultimately resulting in new courses, such as Explaining Physics in physics teacher education, or in the creation of a new teach—study center at one university.

In summary, by the end of the first funding phase, there were theoretically sound, empirically tested models of generic competencies and domain-specific competencies of students in various study domains and corresponding test instruments that had been field tested and found suitable for assessment, as indicated by the results of the validation analyses. However, the results from the first KoKoHs funding phase also indicated that some models and instruments needed further development and validation during the next research stage (2016–2020).

### **Modeling and Measuring Competencies in Higher Education: Validation and Methodological Innovations (KoKoHs II)**

#### *Research Focus, Structure, and Methodological Framework*

During the first funding phase from 2010 to 2015, focus was on conducting fundamental research and developing theory-driven models of generic and domain-specific competencies as well as corresponding assessment instruments. To build on the positive results of the first phase (after a positive external evaluation), to address remaining challenges, and to add a temporal perspective to competency measurement, a new KoKoHs funding phase was launched in 2016 titled Modeling and Measuring Competencies in Higher Education: Validation and Methodological Innovations. The main aim of the new funding phase from 2016 to 2020 is to develop further and validate in depth existing competency models and instruments. To this end, KoKoHs research teams draw on preliminary work, including theory-based models of the competencies to be assessed and studies evidencing the psychometric properties of the instruments. The aim of some of the projects also is to transfer, adapt, and validate empirically tested assessments from

TABLE 2  
*Geographical Range of KoKoHs Assessments according to Clusters*

Range	Generic competencies cluster	Teacher training STEM cluster	Educational science cluster	Economics and social sciences cluster	Engineering cluster
Local assessment	3	4	2	1	
Large-scale assessment	3	3	4	2	3
International assessment				1	

TABLE 3  
*Samples of Field Studies Conducted by the KoKoHs Project Teams*

Cluster	Overall sample sizes	Fields of study	Number of higher education institutions assessed
Generic competencies	7,731	Economic studies, teacher training, psychology, electrical engineering, educational sciences, medicine, sociology, political science	26
Teacher training STEM	12,227	Mathematics, biology, chemistry, physics, information technology	121
Educational science	16,540	Educational sciences, German studies, teaching degrees, history	370
Economics and social sciences	11,858	Business and economic competencies, accounting, entrepreneurship	85
Engineering	11,670	Engineering mechanics, mathematical knowledge in engineering, machine engineering	29

other domains that are more advanced in competency research such as teacher education (see, e.g., the ELMaWi project: *Assessing Subject-Specific Competencies in Teacher Education in Mathematics and Economics*, a quasi-experimental validation study with a focus on domain specificity). The use of empirically proven instruments in new domains is intended to stimulate greater competency orientation in the new domain while keeping development of competency models and instruments efficient.

Compared to assessments in other countries, especially in the United States, the range of assessment approaches in Germany is still rather limited (see Pant & Zlatkin-Troitschanskaia, 2016). Therefore, in some KoKoHs projects, focus is on methodology, including the development and testing of innovative types of assessments of either generic or domain-specific competencies, for example, on the specific consequences of using different item formats (see, e.g., the MultiTex project: *Process-Based Assessment of Multiple Documents Comprehension*). In other projects, KoKoHs researchers undertake comprehensive evaluations of new measurement processes, techniques and tools, innovative assessment designs, or testing practices in higher education, including for example, scoring, validity, and test and scale reliability (for an overview of the KoKoHs projects in the second funding phase, see Pant et al., 2016).

In the KoKoHs research program, focus is on abilities and skills that are (assumed to be) changeable; therefore, in most of the projects, acquisition of and change in competencies as well as how they are influenced by intervention and teaching are investigated (for an example of a project, see the WiWiKom II project in the following: *Valid Assessment of Students' Development of Professional Business and Economic Competencies Over the Course of their Studies: A Quasi-Experimental Longitudinal Study*). While most of the projects during this new KoKoHs phase involve assessment primarily of cognitive aspects of study performance, volitional and motivational facets of competency and the relationship among various competency dimensions also are taken into account.

The new KoKoHs program has a similar structure as the first one (for more details, see Zlatkin-Troitschanskaia, Pant, Kuhn, Toepper, & Lautenbach, 2016). The new program comprises 15 collaborative projects in which generic competencies and domain-specific competencies in teacher education, medicine, economics, and social sciences are assessed. Assessments of generic competencies are tested for their relevance, generality, and applicability across several study domains (e.g., education, psychology, economics, etc.); assessments of domain-specific competencies represent a major part of the core curriculum of related degree

courses. They cover enough content and subject matter to ensure sufficient generality of the skills and abilities in the respective domain examined. In all current KoKoHs projects, undergraduate or graduate degree courses are being examined, and more than one institution of higher education is being assessed in order to provide results that are generalizable beyond individual institutions.

During the current funding phase, the aim of most KoKoHs projects is to analyze the acquisition and change of competencies over time. Two thirds of the projects have longitudinal study designs, and in those projects, competency development in terms of change trajectories over the course of studies is examined. By controlling for person-related and study-related influence factors, the project teams gather evidence of important influences to determine, for instance, the effects of certain learning opportunities on students' acquisition of competencies. Findings on the influences of newly established courses on students' development of competencies have been particularly interesting to educators (for more details, see Pant, Zlatkin-Troitschanskaia, Lautenbach, Toepper, & Molerov, 2016). Innovative formats of competency assessment, such as computer-based and video-based items, are being developed and tested. Computer-based testing and scoring are particularly promising for widespread use of new adaptive assessments in higher education. Scientifically innovative questions concerning competency acquisition and its influences relevant for practice (and society) are being examined using complex study designs and taking various approaches to analysis including experimental or quasi-experimental longitudinal designs and multi-level analyses (for more details, see Pant et al., 2016).

In addition to the analyses of objectivity, reliability, and validity completed in part in preliminary studies, the KoKoHs project teams determine the suitability of assessments by examining especially their convergent, discriminant, incremental, and predictive validity. Through in-depth validation, new assessment and analysis methods, and national and, in some KoKoHs projects, international triangulation of data, the generalizability of results and explanatory power of existing assessments is enhanced to cover new fields of study, different groups of students, and longer periods of time. The performance of specific groups of students of special interest, such as students with a migration background (including refugees), is analyzed in greater depth. Moreover, KoKoHs researchers explore the relationship between changes in generic and domain-specific competencies and ways to foster development of different competency facets through optimized learning opportunities (see, e.g., the ASTRALITE project: Assessment and Training of Scientific Literacy; for more details, see Pant et al., 2016).

Substantial research progress requires the research and results of individual projects to be integrated systematically so as to make them more visible and compatible with national and international research. To this end, overarching

scientific work and meta-studies are being conducted by the Scientific Transfer Project of KoKoHs at Humboldt University of Berlin and Johannes Gutenberg University Mainz to aggregate the findings and results from individual projects and transfer them at various levels to higher education research, practice, and policy. The stated aims and activities of the Scientific Transfer Project serve to channel the overall research output of the KoKoHs program, increase internal and external cooperation, optimize dissemination of results and positioning of KoKoHs research within the German and international research communities, and support the implementation of research results and findings in higher education practice and policy.

#### *Project Example: WiWiKom*

The WiWiKom project (Modeling and Measuring Competencies in Business and Economics Among Students and Graduates by Adapting and Further Developing Existing American and Latin-American Measuring Instruments) and the follow-up WiWiKom II project (Valid Assessment of Students' Development of Professional Business and Economic Competencies Over the Course of Their Studies: A Quasi-Experimental Longitudinal Study) are given as examples to illustrate the conceptual and methodological framework implemented in the KoKoHs projects and highlight how psychometric quality of the assessments was established.

In the WiWiKom projects, focus is on modeling and measuring competencies of students and graduates of business and economics, disciplines in which there was no German language instrument available for reliable and valid assessment of competencies in higher education (Zlatkin-Troitschanskaia, Förster, et al., 2014). To enable empirically based modeling of the levels and structure of business and economic competencies, a domain-specific competency model for business and economics was conceptualized, developed, empirically tested, and validated in German higher education (see Figure 1). Two internationally tested and approved test instruments, the Mexican Examen General para el Egreso de la Licenciatura (EGEL) by the Centro Nacional de Evaluación para la Educación Superior (CENEVAL) and the American Test of Understanding in College Economics (TUCE IV) by the Council for Economic Education (CEE), were adapted, further developed, combined into one German measuring instrument, and comprehensively validated according to the Standards (AERA et al., 2014; Zlatkin-Troitschanskaia, Förster, et al., 2014). Using one German adaptation of the Mexican test and the American test paved the way for comparisons between countries in which the same tests were administered (see, e.g., Brückner, Förster, Zlatkin-Troitschanskaia, Happ, et al., 2015; Brückner, Förster, Zlatkin-Troitschanskaia, & Walstad, 2015).

As a first step, a competency model was developed and validated with respect to content and curricula in higher

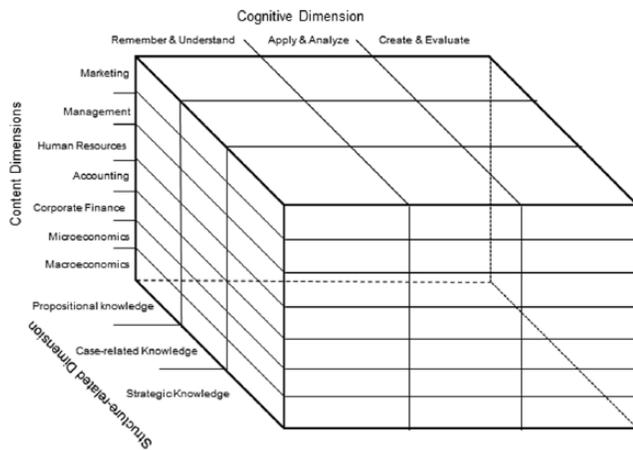


FIGURE 1. Framework model of business and economic competency in the WiWiKom project.

education (see Figure 1). The construct of competency in business and economics was defined in a theory-driven competency model (see Brückner et al., 2014) based on Weinert (2001) and Shavelson (2013) while the assessment design followed Kane’s (2013) interpretative use argument (see also “reasoning from evidence,” Mislevy 1994), the Assessment Triangle by Pellegrino et al. (2001), and the evidence-centered assessment design by Mislevy and Haertel (2006) and Hattie, Jaeger, and Bond (1999). The model differentiates seven domain-specific content dimensions and three levels of cognitive requirements (see Figure 1, Zlatkin-Troitschanskaia, Förster, et al., 2014). The content dimensions represent the core curriculum in business and economics, divided into content areas (e.g., microeconomics, finance, etc.), and the cognitive dimension specifies levels of competency defined in terms of the mental processes (e.g., understanding, applying, etc.) necessary to respond appropriately to cognitive requirements of increasing complexity in professional situations (Zlatkin-Troitschanskaia, Förster, et al., 2014).

Based on this competency model, a test instrument was developed by adapting the EGEL and the TUCE IV. Interviews with experts ( $N = 32$ ) and online ratings ( $N = 78$ ) by lecturers of business and economics were conducted for curricular and content validation and analyses of curricula and textbooks (from 98 degree courses). Students were interviewed in cognitive labs ( $N = 32$ ) to examine their mental processes in order to validate the deployed test items with evidence from the target group of the test and thereby ensure cognitive validity.

All of these steps contributed to a well-founded selection of items, which was thus based on multiple quantitative statistical criteria (e.g., item difficulty) and qualitative content-related criteria (e.g., student performance in the cognitive interviews or judgment by the experts). In total, 220 of the initial 402 items were adapted successfully this way and

were subsumed in 43 test booklets in the form of several nested Youden square designs (Frey, Hartig, & Rupp, 2009). The test was then deployed in the three field surveys at German higher education institutions ( $N = 11,000$  students; 57 universities and colleges) for test calibration, standardization, and establishment of validity of the internal structure (Zlatkin-Troitschanskaia, Förster, et al., 2014). Analysis of the data collected indicated, for instance, that students in business and economics show very diverse preconditions at the beginning of their studies. A wide range of competency levels (Brückner, Förster, Zlatkin-Troitschanskaia, Happ, et al., 2015) and considerable gender differences in academic performance resulting from commonly employed teaching and examination methods (Brückner, Förster, Zlatkin-Troitschanskaia, & Walstad, 2015) were found.

In the WiWiKom II project, which is based on the competency model developed and tested and the test instrument used in WiWiKom I, more in-depth validation questions are addressed, and the results so far are being broadened by including individual change measurement (Zlatkin-Troitschanskaia, Pant, Förster, Brückner, & Fox, 2016). The project follows a quasi-experimental design, examining competency constructs and generic cognitive abilities according to the multitrait-multimethod matrix approach and two-group comparisons (students in social sciences and in economics) for convergent, discriminant, incremental, and predictive validation. Quasi-experimental variation is achieved through sampling that includes participants not only in two study domains (social sciences and business and economics) but also in different stages of study programs (orientation, advanced, and specialization study phases in the bachelor degree courses and transition into master degree courses or the job market) at 20 universities in Germany.

In WiWiKom II, business and economic competencies and their development over bachelor degree programs are modeled as central dependent variables. In a longitudinal study, the extent to which these dependent variables can be explained by variables related to business and economics degree courses (e.g., learning opportunities attended over the course of bachelor studies) is examined. In further validation analyses, focus is on differentiation in measurement between the dependent variables and theoretically related criteria (e.g., school grade point average and generic cognitive abilities) and correlations between the dependent variables and construct-relevant external criteria (e.g., students’ final grades for their bachelor degree). In a quasi-experimental comparative analysis, further aspects of discriminant validity are examined by analyzing the manifestation of the dependent variables within the target group (students in business and economics) and a control group (students in social sciences) in cross-sectional and longitudinal studies (for more detail, see Zlatkin-Troitschanskaia, Pant, Förster, et al., 2016).

### **Conclusion: Lessons Learned, Challenges, and Further Perspectives**

Over the past two decades, higher education in Germany (and Europe) has seen far-reaching educational reforms, following the Bologna reform and implementation of the New Governance Model. Comprehensive restructuring processes have been initiated to address inherent problems in higher education. In recent years, some of these problems have deteriorated and new challenges have arisen, including considerable social pre-selection in the new bachelor and master degree programs, unequal access to higher education for students with a migration background, gender inequality in various disciplines, high dropout rates, and long study durations (OECD, 2015). The quality, effectiveness, efficiency, and individual and societal returns of higher education currently are being given increasing attention. Many decisions related to these issues should be informed by empirical data on the influences, development, design, and effects of academic teaching-and-learning processes (see, e.g., Coates, 2016). To this end, intended outcomes of academic learning processes, such as the competencies students are to acquire in higher education, need to be defined, modeled, and assessed in a reliable and valid way.

Both challenges and promising perspectives lie in the development and testing of new assessments for higher education (see also Shavelson et al., 2015). Modern higher education institutions that truly embrace and tap the potentials of diversity need assessments comprising realistic, action-oriented items that enable institutions to evaluate the state of learning of their students in line with the demands of society in the 21st century. Such assessments can be developed only in cooperation among experts of content, teaching, and measurement and require very thorough and comprehensive testing before they can be used in practice.

Using the assessments developed in the KoKoHs research program in Germany, the aforementioned challenges of higher education can be described in greater detail than before at the individual and institution levels. By controlling for suitable influence factors, important determinants of learning outcomes can be identified, and targeted improvement measures can be devised.

One lesson learned from KoKoHs is that for substantial research progress to be made, the results of individual projects must be integrated systematically and made more visible and compatible with national and international research. In KoKoHs, a number of measures served these purposes, including establishing a dedicated coordination and transfer project, publishing meta-analyses and program overviews, implementing an explicit overall program strategy and common assessment design frameworks, conducting regular round tables to discuss and agree on approaches to relevant research challenges, establishing and maintaining cooperation with leading international colleagues in the different

fields, and organizing conferences as well as symposia at international conferences such as the AERA annual meeting.

In KoKoHs, interdisciplinarity from the outset was an important prerequisite of the program, and the collaboration among and between teams, always including experts in subject matter and methodology as well as top-level international researchers, provided appropriate expertise and supported conceptual specification of domain models, test operationalizations, measurement models, validation analysis, and overall research strategy. The experience from KoKoHs shows that despite high-level conceptual and methodological requirements, students' learning outcomes in higher education can be assessed in an objective, reliable, and valid way. High-quality assessments can be developed and used in higher education by following systematic approaches centered on quality standards of testing (AERA et al., 2014).

At the beginning of the program, a new research community had to be established that could undertake the necessary long-term research in the field of empirical educational research in Germany. The promotion of young researchers is one of the key objectives stated, which has been pursued by organizing interdisciplinary colloquia, methodology workshops on many relevant topics from item design to cognitive validation, and attractive mentoring and presentation opportunities for outstanding young researchers.

The KoKoHs experience has shown that, with such complex projects, awareness of coordination efforts and needs, as with various other components, is not always a given during the planning and application stages of the project and needs to be promoted while additional funding may need to be attracted over the course of the project. In the case of the KoKoHs program, fortunately, project teams could apply for additional purposeful funding from the BMBF, for example, for extended documentation and data management efforts. With varying project runtimes, maintaining the knowledge and expertise acquired during the first phase has been a particular challenge. Results were usually published at the end of the project runtimes or even beyond; a large proportion of the project teams from the first phase were not granted funding for follow-up projects in the new phase. Hence, integration of findings and inclusion of assessment expertise from the first phase has been a particular challenge for the Scientific Transfer Project (see an overview in Zlatkin-Troitschanskaia et al., 2017).

In the KoKoHs program, suitable competency models and corresponding assessment instruments were developed and comprehensively validated for a number of major disciplines and fundamental student competencies. Developing a common conceptual understanding and methodological standards has been key from the outset. The publications of the individual projects offer an orientation for the specific domains, and some of the assessments can be adapted to

other domains, which is an important focus of the second KoKoHs phase. With regard to adaptations of instruments from or into other languages and cultural contexts, the experience from international adaptations in KoKoHs highlighted feasibility of international comparative studies based on similar curricula and additional comparability analyses. Comparative findings have helped identify international best practice examples in higher education; for example, one of the countries assessed did not show gender effects in economic knowledge, which are a recurring problem in other countries (Brückner, Förster, Zlatkin-Troitschanskaia, & Walstad, 2015). Nevertheless, the experiences made in KoKoHs showed that the intention to adapt tests or generalize interpretations need to be taken into account during the test development and validation stages (Zlatkin-Troitschanskaia, Kuhn, et al., 2014).

Particularly with regard to the use of test instruments, assessment experts highlighted the importance of documenting and communicating very explicitly the constructs being assessed, the possible uses for and interpretations of which the assessments have been validated. While research teams publish limitations of interpretation in their validation studies, many test users do not have the psychometric expertise to evaluate these studies and require clear instruction in the test documentation and interpretation aids (Koretz, 2016). To promote the transfer of assessments into practice, KoKoHs teams have devoted additional attention to develop user-oriented feedback systems, aids to interpretation of assessment results, and teaching-and-learning tools for students, teachers, and university management. In this, KoKoHs researchers strive to identify best practices and examine the conditions for successful implementation in practice.

Transfer of assessments into practice is a key driver of improvement toward more competency-oriented learning and teaching. Transfer activities have involved institutions that participated in initial validation assessments and continue to use the assessments on a regular basis. Maintaining positive relationships with higher education institutions has been an important goal; trust has been built by designing assessments specifically for improvement and explicitly not for simplified rankings. Institutions have adopted findings on students' competency levels to design new competency-oriented courses, degree programs, curricula, and instruments to be implemented as examinations, for evaluation of courses, accreditation of programs, or as a basis for a new competency-oriented graduate school at one of the participating universities. The KoKoHs program has contributed to improving teaching and examination practice in higher education in Germany by highlighting successful approaches and revealing huge deficits in student competencies, and it has contributed to advancing current competency research internationally by providing orientation for the implementation of similar projects in other countries (for more details, see Zlatkin-Troitschanskaia, 2017).

## References

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2004). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Biggs, J., & Tang, C. (2011). *Teaching for quality learning at university*. Maidenhead, UK: McGraw-Hill and Open University Press.
- Brückner, S., Förster, M., Zlatkin-Troitschanskaia, O., Happ, R., Walstad, W. B., Yamaoka, M., & Asano, T. (2015). Gender effects in assessment of economic knowledge and understanding: Differences among undergraduate business and economics students in Germany, Japan, and the United States. *Peabody Journal of Education, 90*(4), 503–518.
- Brückner, S., Förster, M., Zlatkin-Troitschanskaia, O., & Walstad, W. B. (2015). Effects of prior economic education, native language, and gender on economic knowledge of first-year students in higher education. A comparative study between Germany and the USA. *Studies in Higher Education, 40*(3), 437–453.
- Brückner, S., Zlatkin-Troitschanskaia, O., & Förster, M. (2014). Relevance of adaptation and validation for international comparative research on competencies in higher education—A methodological overview and example from an international comparative project within the KoKoHs research program. In F. Musekamp, & G. Spöttl (Eds.), *Competence in higher education and the working environment. National and international approaches for assessing engineering competence. Vocational education and training: Research and practice* (Vol. 12, pp. 133–152). Frankfurt: Peter Lang.
- Coates, H. (2014). (Ed.). *Higher education learning outcomes assessment: International perspectives*. Frankfurt: Peter Lang.
- Coates, H. (2016). Assessing student learning outcomes internationally: Insights and frontiers. *Assessment and Evaluation in Higher Education, 41*, 662–676. doi:10.1080/02602938.2016.1160273
- European Commission. (2014). *Interim evaluation of the strategic framework for European cooperation in education and training (ET 2020)*. Luxembourg: Author.
- Ewell, P. T. (2005). Can assessment serve accountability? It depends on the question. In J. C. Burke & Associates (Eds.), *Achieving accountability in higher education* (pp. 1–24). San Francisco, CA: Jossey-Bass.
- Frey, A., Hartig, J., & Rupp, A. A. (2009). An NCME instructional module on booklet designs in large-scale assessments of student achievement: Theory and practice. *Educational Measurement: Issues and Practice, 28*(3), 39–53.
- Groß Ophoff, J., Schladitz, S., Leuders, J., Leuders, T., & Wirtz, M. (2015). Assessing the development of educational research literacy. The effect of courses on research methods in studies of educational science. *Peabody Journal of Education, 90*, 560–573.
- Hartmann, S., Upmeier zu Belzen, A., Krüger, D., & Pant, H. A. (2015). Scientific reasoning in higher education: Constructing

- and evaluating the criterion-related validity of an assessment of preservice science teachers' competencies. *Zeitschrift für Psychologie*, 223, 47–53.
- Hattie, J., Jaeger, R. M., & Bond, L. (1999). Persistent methodological questions in educational testing. In A. Iran-Nejad, & P. D. Pearson (Eds.), *Review of research in education* (pp. 393–446). Washington, DC: American Educational Research Association.
- Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, 50(1), 1–73.
- Koretz, D. (2016, April). *Making use of competency assessments in higher education: Lessons from K–12 testing*. Paper presented at the Kick-Off-Meeting of the Research Program Modeling and Measuring Competencies in Higher Education—Validation and Methodological Innovation, Berlin.
- Kuhn, C., Toepper, M., & Zlatkin-Troitschanskaia, O. (2014). *Current international state and future perspectives on competence assessment in higher education—Report from the KoKoHs affiliated group meeting at the AERA Conference on April 4, 2014 in Philadelphia (USA)* (KoKoHs Working Papers, 6). Berlin & Mainz: Humboldt University & Johannes Gutenberg University.
- Kultusministerkonferenz [Standing Conference of the Ministers of Education and Cultural Affairs of the Federal Republic of Germany]. (2000). *Rahmenvorgaben für die Einführung von Leistungspunktsystemen und die Modularisierung von Studiengängen* [Framework for the introduction of credit point systems and the modularization of degree courses.]. Retrieved May 26, 2016, from <http://www.kmk.org/fileadmin/Dateien/pdf/PresseUndAktuelles/2000/module.pdf>
- Leutner, D., Fleischer, J., Grünkorn, J., & Klieme, E. (Eds.). (in press). *Competence assessment in education: Research, models and instruments*. Heidelberg: Springer.
- Mislevy, R. (1994). Evidence and interference in educational assessment. *Psychometrika*, 59, 439–483.
- Mislevy, R., & Haertel, G. D. (2006). Implications of evidence-centered design for educational testing. *Educational Measurement*, 25(4), 6–20.
- Musekamp, F., Spöttl, G., Mehrafza, M., Heine, J.-H., & Heene, M. (2014). Modeling of competences for students of engineering mechanics. *International Journal of Engineering Pedagogy*, 4(1), 4–12.
- Obersteiner, A., Moll, G., Reiss, K., & Pant, H. A. (2015). Whole number arithmetic—Competency models and individual development. In X. Sun, B. Kaur, & J. Novotná (Eds.), *Proceedings of the Twenty-Third ICMI Study: Primary mathematics study on whole numbers* (pp. 235–242). Macao: University of Macau.
- OECD. (2015). *Education at a glance 2015: OECD indicators*. Paris: Author.
- Pant, H. A., & Zlatkin-Troitschanskaia, O. (2016). Valid assessment of student competencies in higher education—Challenges, methodological innovations, and perspectives for educational measurement [Special issue]. *Journal of Educational Measurement*, 53(3).
- Pant, H. A., Zlatkin-Troitschanskaia, O., Lautenbach, C., Toepper, M., & Molerov, D. (Eds.). (2016). *Modeling and measuring competencies in higher education—Validation and methodological innovations* (KoKoHs Working Papers 11). Berlin & Mainz: Humboldt University & Johannes Gutenberg University.
- Pellegrino, J. W., Chudowsky, N., & Glaser, R. (Eds.). (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Academies Press.
- Rychen, D. S. (2004). Key competencies for all: An overarching conceptual frame of reference. In D. S. Rychen, & A. Tiana (Eds.), *Developing key competencies in education: Some lessons from international and national experience* (pp. 5–349). Paris: UNESCO.
- Schneider, C. G. (2009). The proof is in the portfolio. *Liberal Education*, 95(1). Retrieved from <https://www.aacu.org/publications-research/periodicals/proof-portfolio>
- Shavelson, R. (2013). On an approach to testing and modeling competence. *Educational Psychologist*, 48(2), 73–86.
- Shavelson, R., Davey, T., Ferrara, S., Holland, P., Webb, N., & Wise, L. (2015). *Psychometric considerations for the next generation of performance assessment*. Princeton, NJ: Educational Testing Service.
- Soares, L. (2012). *A “disruptive” look at competency-based education. How the innovative use of technology will transform the college experience*. Retrieved from [https://cdn.american-progress.org/wp-content/uploads/issues/2012/06/pdf/comp\\_based\\_education.pdf](https://cdn.american-progress.org/wp-content/uploads/issues/2012/06/pdf/comp_based_education.pdf)
- Tiffin-Richards, S. P., & Pant, H. A. (in press). Arguing validity in educational assessment. In D. Leutner, J. Fleischer, J. Grünkorn, & E. Klieme (Eds.), *Competence assessment in education: Research, models and instruments*. Heidelberg: Springer.
- Toepper, M., Zlatkin-Troitschanskaia, O., Kuhn, C., Schmidt, S., & Brückner, S. (2014). *Advancement of young researchers in the field of academic competency assessment—Report from the International Colloquium for Young Researchers from November 14–16, 2013 in Mainz* (KoKoHs Working Papers 5). Berlin & Mainz: Humboldt Universität & Johannes Gutenberg University.
- Weinert, F. E. (2001). Concept of competence: A conceptual clarification. In D. S. Rychen, & L. H. Salganik (Eds.), *Defining and selecting key competencies* (pp. 45–65). Seattle, WA: Hogrefe and Huber.
- Wissenschaftsrat. (2015). *Empfehlungen zum Verhältnis von Hochschulbildung und Arbeitsmarkt. Positionspapier vom 16.10.2015 (Drs. 4925-15)* [Recommendations on the relationship between higher education and the job market. Position paper, 16 Oct 2015]. Bielefeld: Author.
- Zlatkin-Troitschanskaia, O., Förster, M., Brückner, S., & Happ, R. (2014). Insights from a German assessment of business and economics competence. In H. Coates (Ed.), *Higher education learning outcomes assessment—International perspectives* (pp. 175–197). Frankfurt: Peter Lang.
- Zlatkin-Troitschanskaia, O., Kuhn, C., & Toepper, M. (2014). Modeling and assessing higher education learning outcomes in Germany. In H. Coates (Ed.), *Higher education learning outcomes assessment—International perspectives* (pp. 213–235). Frankfurt: Peter Lang.
- Zlatkin-Troitschanskaia, O., Pant, H. A., Förster, M., Brückner, S., & Fox, J.-P. (2016). WiWiKom II—Valid assessment of students' development of professional business and economic competencies over the course of their studies—A quasi-experimental longitudinal study. In H. A. Pant, O. Zlatkin-Troitschanskaia, C. Lautenbach, M. Toepper, & D. Molerov (Eds.), *Modeling and measuring competencies in higher education—Validation*

and methodological innovations (KoKoHs) (pp. 57–60). Berlin & Mainz: Humboldt Universitaet & Johannes Gutenberg University.

Zlatkin-Troitschanskaia, O., Pant, H. A., Kuhn, C., Toepper, M., & Lautenbach, C. (2016). Assessment practices in higher education and results of the German research program modeling and measuring competencies in higher education. *Journal of Research & Practice in Assessment, 11*(1), 46–54.

Zlatkin-Troitschanskaia, O., Pant, H. A., Lautenbach, C., Molerov, D., Toepper, M., & S. Brückner (2017). *Modeling and measuring competencies in higher education. Approaches to challenges in higher education policy and practice*. Heidelberg: Springer.

Zlatkin-Troitschanskaia, O., & Shavelson, R. (2015). Competence assessment in higher education [Special issue]. *Studies in Higher Education, 40*(3).

Zlatkin-Troitschanskaia, O., Shavelson, R. J., & Kuhn, C. (2015). The international state of research on measurement

of competency in higher education. *Studies in Higher Education, 40*(3), 393–411.

### Authors

OLGA ZLATKIN-TROITSCHANSKAIA is chair of business and economics education at Johannes-Gutenberg University Mainz, Germany.

HANS ANAND PANT is chair of research methods in education at Humboldt Universitaet Berlin, Germany.

MIRIAM TOEPPER is research assistant in the Department of Business and Economics Education at Johannes Gutenberg University Mainz.

CORINNA LAUTENBACH and DIMITAR MOLEROV are research assistants in the Department of Education Studies at Humboldt Universitaet Berlin, Germany.