

RESEARCH

Open Access

# Assessing the financial knowledge of university students in Germany

Manuel Förster\*, Sebastian Brückner and Olga Zlatkin-Troitschanskaia

\* Correspondence:  
manuel.foerster@uni-mainz.de  
Johannes Gutenberg University, Mainz  
Institute of Business and Economics  
Education, Jakob-Welder-Weg 9,  
D-55099 Mainz, Germany

## Abstract

**Background:** The importance of professional competence (PC) in business administration (BA) has increased considerably in many industrial nations over the past several years. However, while economic competence is being assessed internationally in the Assessment of Higher Learning Outcomes (AHELO) study, the modeling and assessment of PC in BA needs more research, particularly from an international perspective. We defined and modeled the construct of PC in BA based on theoretical analyses and evidence from international studies and, for the assessment, focused on knowledge of BA as a key facet of PC in BA. In this article, we describe the developed structural model of knowledge in BA and present the specifications and findings for the example of financial knowledge (FK). In the model, we describe cognitive levels of FK in relation to subject content and subject didactics. Moreover, we discuss influence factors on FK.

**Method:** Subsequently, we present the results from the empirical analyses on FK. Assessment data was gathered using an adapted and further developed international test instrument. The sub-sample for the analyses comprised 773 students from 23 institutions of higher education in Germany. We used item response models to confirm the theoretically modeled levels and multilevel modeling to analyze influence factors on FK.

**Results:** The Rasch model showed a good fit to the data and confirmed the theoretically modeled levels. From the perspective of vocational education and training, we investigated the extent to which FK is influenced positively by commercial vocational training completed prior to higher education studies. We analyzed this while controlling personal influence factors, such as mother tongue and gender, and study-related influence factors such as completion of subject-related courses at university, number of semesters, and type of institution of higher education. We found that prior commercial vocational training affected FK even when the other influence factors were controlled.

**Conclusions:** These results support the assumption that during dual vocational education and training students acquire professional knowledge and gain experience related to their job or practical training that are not or cannot be taught in this way at universities or universities of applied sciences.

**Keywords:** Professional competences in finance; Measuring/assessing knowledge; Item response theory models (IRT); Multilevel models (HLM)

## Background

### State of research

The importance of professional competence (PC) in business administration (BA) has increased considerably in many industrial nations over the past several years (Ministerial Council for Education and Early Childhood Development and Youth Affairs MCEECDYA 2011). Knowledge of business content is crucial to the success of companies. However, while economic competence is being assessed internationally in the Assessment of Higher Learning Outcomes (AHELO) study (Tremblay et al. 2012), the modeling and assessment of PC in BA needs more research, particularly from an international perspective. Preliminary approaches in this area focused on commercial competence in pre-university education, particularly in commercial vocational education and training (VET) (e.g., Mentele et al. 2013). However, it is usually the graduates of universities and universities of applied sciences who take key positions in business management and are responsible for the success of companies.<sup>a</sup>

In the context of the Bologna reform, and with increasing competition due to internationalization, the kinds and structures of study programs in BA in higher education have diversified enormously. For companies and the economy, it is increasingly important to know what PC students acquire in BA at institutions of higher education. For example, some business representatives express great skepticism with regard to the quality of the new bachelor level degree programs in business administration. Small and medium-sized businesses particularly distrust the quality of the new bachelor level degree programs (see Jahn 2007) and, because there is little empirical data, it is a huge challenge for companies to judge the PC of bachelor level degree students in BA.

From a corporate perspective, it is particularly important to assess student and graduate knowledge of BA, along with social and motivational aspects (Zlatkin-Troitschanskaia et al. 2014). Knowledge of BA is an essential condition for developing PC in BA (Größler et al. 2002). In Germany, responsibility for teaching business content lies mainly within universities and universities of applied sciences, the two major types of higher education institutions.<sup>b</sup> The structures of study programs have been subject to harmonization processes for years (e.g., Krücken 2004), for example, through uniform designation of degrees, such as the Bachelor of Arts and Master of Arts. Nevertheless, the public's perception is that the quality of content taught and degrees to be obtained differ between institutions of higher education (e.g., Nickel 2011). Company administrators need to know whether differences exist, what kind of differences to expect with regard to graduate knowledge of BA and whether they need to take into account differences between educational institutions, such as between universities and universities of applied sciences.

Some evidence of differences in knowledge among students even from *identically labeled* degree courses was provided by the Innovative Teach-Study Network in Academic Higher Education (ILLEV) study (e.g., Happ et al. 2013). In this study, students' declarative knowledge was assessed in a number of subdomains of BA. The results showed that students from *identically labeled* business and economic degree courses (e.g., M.Sc. or M.A.) differed greatly in their business knowledge even if the assessment was controlled for the courses the students attended. The findings illustrate that degree labels and certificates provide little indication of the type, extent, and quality of knowledge students acquire. In the ILLEV project, the Business Administration Knowledge

Test (BAKT, Bothe T, Wilhelm O, Beck K (2006): Business administration knowledge. Assessment of declarative business administration knowledge: Measurement development and validation, unpublished manuscript.) was administered, as it is generally suitable for assessing knowledge of BA. However, a major limitation of the BAKT is that it assesses declarative knowledge only (Zlatkin-Troitschanskaia et al. 2014).

Comprehensive modeling and a valid assessment of business knowledge requires first, an in-depth analysis of the practical requirements of business professions and, second, a test instrument that adequately represents these professional requirements. This complex task is being undertaken in Project WiwiKom,<sup>c</sup> which assesses PC in BA. The project team defined and modeled the construct of PC in BA based on theoretical analyses and evidence from international studies and, for the assessment, the team focused on business knowledge as a key facet of PC in BA. To enable international comparability of results, which is relevant for knowledge of BA, the project team took an internationally tested instrument from Mexico and developed it further in accordance with the model. This test instrument was used to assess students in a pilot study in 2012 and in two large-scale surveys at 33 institutions of higher education in Germany in 2013.

We present analyses of student knowledge of BA based on data from the first large-scale survey and then identify the extent to which it differs among various institutions and types of institutions of higher education in Germany. For conciseness, we focus on the findings in financial knowledge (FK) as an example of a key facet of knowledge of BA (Porter 2000). Moreover, textbook analyses identified the financial sector as central economic content (Lauterbach 2013) and, in the area of vocational schools, FK was also classified as a central component of business knowledge (Preiss 2005). In the article, we describe the structural model of knowledge of BA as developed in Project WiwiKom and the specifications for the example of FK. With regard to the model, we describe cognitive levels of FK in connection with subject content and subject didactics, as well as how they were represented by the items. Based on Project WiwiKom data, we present the results of the empirical analyses with regard to the following questions. First, do we have empirical evidence of the modeled cognitive levels in terms of gradual differences in FK for the entire construct or one or several content subdimensions? Second, which factors influence FK? From the perspective of commercial VET, it would be most relevant to know the extent to which FK is influenced positively by commercial VET completed prior to studies in higher education. We analyzed this while controlling for personal influence factors, such as gender and mother tongue, and study-related influence factors, such as completion of subject-related courses at university, number of semesters, and type of institution of higher education.

### **Financial knowledge as a key facet of professional competence in business administration**

Following the state of international research, we understand PC in BA as the cognitive disposition necessary for successful decision-making in professional situations within a company (see in detail Zlatkin-Troitschanskaia et al. 2014). In this, we focused on knowledge of BA. A knowledge construct can be specified with regard to content dimensions and cognitive dimensions (e.g., Alexander et al. 1994).

Identifying appropriate content dimensions of the construct is crucial because decision-making in practical situations in a company requires knowledge of many content areas. In Project X, we differentiated the business content dimensions of human resources, finance, accounting, marketing, and organization and management, in line with common classifications of the domain of BA in the literature and at institutions of higher education in Germany (see in detail Zlatkin-Troitschanskaia et al. 2014). Since many companies are organized functionally, the above distinction of content areas corresponds with common practical distinctions of divisions or departments in a company. From the company perspective, it is important that students acquire both general and area-specific knowledge so that they can both understand the company in its entirety and also work professionally in specific departments. Companies can be regarded as information systems (Preiss 2005), in which the area of finance is a key integrative component.

With regard to the cognitive dimensions of knowledge of BA, the project team made assumptions about general levels that are applicable to all content dimensions of BA. Following Anderson and Krathwohl (2001) and Walstad et al. (2007), the following three hierarchical cognitive process levels were differentiated: (1) remembering and understanding, (2) applying and analyzing, and (3) creating and evaluating (for further information about content-based modeling of FK see Zlatkin-Troitschanskaia et al. 2014). These rather abstract theoretical levels were specified using item difficulty generating characteristics (Hartig and Frey 2012) for each content dimension of BA. In the following, we illustrate the conceptualization of these characteristics for the example of the content dimension of finance.

FK is a key dimension of knowledge of BA. It is necessary for making professional financial decisions in a company and solving financial problems. FK serves essentially to control decision processes by providing decision-makers in companies with information about finance and cash flows as a basis for informed investment decisions. Financial decisions serve to maintain the financial stability of a company while other key objectives, such as profitability, are being pursued (Becker 2012). Accordingly, FK includes knowledge of financial concepts and how to handle risks. With internationalization, increasing competitive pressure, the shortening of product life cycles, and many other developments, financial decisions are gaining increasing importance for a company's ability to survive (Prätsch et al. 2003).

For the modeling, the project team conducted subject-didactic based content analyses, from textbooks and curricula for BA in higher education and commercial VET, an online rating (N = 78), expert interviews with professors and lecturers (N = 32) and cognitive interviews (N = 20) to determine item difficulty generating characteristics. Thereby, a selection of textbooks, which are most frequently used in financial studies, were referred to in order to ensure the conformity of the following steps to the contents in bachelor level degree studies. For this purpose, subject-related didactical conceptions of textbooks, which focus on sequences of teaching contents, were taken as a basis for the definition of complexity characteristics. Lecturers in the respective subject areas assessed the level of difficulty of each item via an online rating and this was discussed with experts afterwards. Here, experts compared the criteria which contribute to a rather difficult or rather easy item in contrast to other items. Within the scope of cognitive interviews with students, the statements of students were analyzed. Thus,

several items were identified as criteria which influence the difficulty level, for example, the kind of mathematical modeling or the elaboration of financial concepts. On the basis of these validation analyses, FK was modeled into four hierarchical levels. The four levels were conceptualized according to increasing difficulty in the financial decision-making process, based on the respective information that needed to be considered and the decision perspectives that needed to be weighed. Thus, differentiation of the levels was based mainly on the following criteria for determining the difficulty of financial decisions and test items:

- 1) period of time to be considered for the decision, for example, taking into account information from one time period as opposed to several time periods;
- 2) the uncertainty to be considered for the decision, for example, using available, factual data only, instead of data based on estimates of future developments;
- 3) the degree of abstraction of the decision, for example, understanding simple data, such as simple inflows or outflows, as opposed to handling information that is condensed in certain key figures or accumulated from various business areas, and
- 4) mathematical and algorithmic modeling, from simple summation to linking different basic arithmetic operations and to operations, which demand the application of complex formulas.

According to these characteristics, we assume that increasing periods of time which are taken into consideration, a higher level of insecurity in a particular situation, a higher degree of abstraction as well as complex requirements for mathematical and algorithmic modeling all lead to more difficulty in making financial decisions. As a result, a variety of combinations of these four criteria is possible. Thereby, the following four are especially relevant to the domain of finance in bachelor level studies (see Table 1).

- 1) Understanding basic financial concepts in a company

At the lowest difficulty level, test-takers need to be able to understand basic financial concepts. To this end, the ability to identify value streams is of central importance, since value streams show condensed financial information and are vital to all parts of a business, as described in Porter's value chain (Porter 2000). Value streams are a primary source of information about the situation of a company and are considered by employees and managers in making responsible financial decisions in a company. Becker's textbook on investment and finance (2012) starts with basic value streams in a company, such as cash inflows and outflows. The textbook explains inflows and outflows as the basis of all budgeting, financial appraisal and control. The textbook by Zantow and Dinauer (2011) starts with an explanation of the differences between terms for value streams in companies, such as cash inflow and revenue. At Level 1 of FK, test-takers need to be able to understand and recall basic financial concepts. Thus, at this level, the basics of investment and finance include identifying, differentiating, structuring, and categorizing value streams. Structuring and categorizing are classified into this category, because a comprehension of single core concepts (without transferring them to new situations) is necessary. To reach Level 1, test-takers do not need more advanced

**Table 1 Overview of levels of financial knowledge, characteristics of difficulty and levels of knowledge**

Levels of FK	Characteristics of difficulty in differentiating the levels
1) Understanding basic financial concepts in a company	<i>Time periods to be considered:</i> past <i>Uncertainty to be considered:</i> no <i>Degree of abstraction:</i> low <i>Calculation:</i> no or simple additive calculation
2) Analyzing (static) concepts and (static) investment appraisal methods	<i>Time periods to be considered:</i> mainly past <i>Uncertainty to be considered:</i> usually no <i>Degree of abstraction:</i> medium <i>Calculation:</i> combination of different basic arithmetic operations
3) Applying and analyzing dynamic investment appraisal methods	<i>Time periods to be considered:</i> several periods of time, usually past and future <i>Uncertainty to be considered:</i> usually yes <i>Degree of abstraction:</i> high <i>Calculation:</i> calculations based on formulae
4) Creating finance plans	<i>Time periods to be considered:</i> several periods of time, past and future <i>Uncertainty to be considered:</i> yes <i>Degree of abstraction:</i> very high <i>Calculation:</i> calculations based on formulae

knowledge of investment and financial instruments or forecasting abilities. At this level, test-takers need to consider data based on past financial events without any uncertainty. There is a low degree of abstraction, since the main concepts to be considered in the decision process or item response process are cash inflows and outflows.

2) Analyzing (static) concepts and (static) investment appraisal methods

In the introduction to their textbook, Prättsch et al. (2003) describe the solving of financial problems as a key operation in a company. Financial information is necessary for all strategic decisions in a company. First, it needs to be identified and then be prepared for further use in decision-making. Test-takers need to analyze single pieces of information and combine and process them. In this context, Becker (2012) introduces the terminology and calculation of profitability indices and includes the analysis of liquidity indices and cash flow values. Zantow and Dinauer (2011) subsume these concepts under financial goals. To evaluate and calculate profitability and liquidity indices, it is necessary to have conceptual knowledge of profits and means of payment as well as be able to use single pieces of information that are financially relevant. Furthermore, test-takers need to be able to understand different terminologies and value streams and to combine numerical values. This is more difficult than the requirements at Level 1. At Level 2, we considered calculations of the cost of capital equally difficult to calculations of liquidity indices, since the calculations of both types of financial items involve several numbers and must both be analyzed. The difference in difficulty can also be explained by the kind of inference the test-taker is using (Minnameier 2013). While the memorization of a more difficult concept can be regarded as an abductive inference,

the calculation itself can be regarded as a more deductive inference, which is a more logical conclusion from a mental representation given, needed at almost all levels of item difficulty. Often the necessary operations include calculating averages, percentages, and other values that are preliminary approximations of forecasting, such as the weighted average cost of capital, which is used for discounting future cash flow. At Level 2, the index calculations include actual figures only. Concepts described at Level 1, such as cash inflows and outflows reappear but are not considered individually and need to be combined in a calculation or other method. At Level 2, the degree of abstraction is higher since inflows and outflows are condensed into new, more complex figures. There is only little uncertainty to be taken into account, since financial concepts are static at this level.

### 3) Applying and analyzing dynamic investment appraisal methods

Apart from statically aggregated indicators, financial decisions also need to reflect time and risk factors. These additional variables lead to an increase in difficulty. The variables of time and risk can be subject to various changes; they are dynamic elements in financial decision-making. In the textbooks analyzed (e.g., Becker 2012; Zantow and Dinauer 2011), investment appraisal methods are introduced only after the sections on identifying value streams and analyzing and evaluating profitability and cash flow. In Project WiwiKom, the test items on investment appraisal methods have higher cognitive requirements for decision-making in finance. The required operations include comprehensive, complex calculations with prognostic, multi-period components. However, in contrast to Level 2, test-takers need to consider data from several periods of time and to handle some uncertainty. Therefore, we assumed Level 3 to have a slightly higher degree of difficulty, which should also be reflected in the data. The content, such as the application of dynamic investment methods is, according to expert opinion, assessed as more difficult than the identification of value flows, as found at Level 1.

### 4) Creating finance plans

Test-takers who successfully understand, apply, and analyze static and dynamic investment appraisal methods are able to use key financial concepts and instruments professionally. In addition, companies need to make decisions for strategic planning. Corporate finance planning requires a systematic, multi-perspective view of finance, taking into account both various types of financial information from individual departments and also an overall understanding of the entire company. The textbook analysis showed clearly that the creation of finance plans requires elaborate FK and an understanding of the entire company (e.g., Becker 2012; Zantow and Dinauer 2011). At Level 4, test-takers need to forecast and evaluate the future development of the company. Level 4 has the highest degree of abstraction, as decision-making is based on different types of information from different departments. In the course of the validation studies, it was obvious that there were no items from the adapted test version which matched to these characteristics. Therefore, these characteristics could not be operationalized. Items at a higher level of difficulty or other item formats need to be newly developed and validated in follow-up studies.

Content analyses indicated that FK is a one-dimensional construct with regard to content. Even though we assumed several levels of difficulty, they all referred in content to financial decision-making.

We empirically examined this theoretical model of FK and the derived levels from multiple evidence bases. In the following, we present the testing instrument, data, methods, and results as to whether the empirical analyses confirmed the above model, whether the modeled difficulty levels of the items are reflected in the item parameter estimations, and if the items are represented in one dimension (Question 1 above).

Furthermore, we present our results concerning the extent to which FK was positively influenced by *commercial vocational training* completed prior to studies in higher education, when personal influence factors, such as mother tongue and gender, and study-related influence factors, such as completion of subject-related courses at university, number of semesters, and type of institution of higher education are controlled (Question 2 above).

Since finance is a central part of companies, it is an important learning domain for commercial trainees. Students who have already completed an apprenticeship should possess previous knowledge which is subject relevant. The positive influence of economic and didactical knowledge has already been established in empirical studies (cf. Kuhn et al. 2014). In the course of this project, a positive effect of vocational education on economic knowledge was confirmed (Brückner et al. 2015; Zlatkin-Troitschanskaia et al. 2015).

## Method

### Testing instrument

Project WiwiKom assessed the FK and the cognitive processes of students in BA degree courses at institutions of higher education in Germany. This group of prospective financial decision-makers was assessed in order to draw conclusions about their understanding of BA concepts in professional situations. The assessment focused on the knowledge that is necessary for making financial decisions in companies. The collected data was meant to provide evidence as to whether or not students acquire sufficient FK during their studies to address appropriate finance issues in a company.

To this end, the project team adapted and developed further the Mexican 'Examen General para el Egreso de la Licenciatura Administración' (Centro Nacional de Evaluación para la Educación Superior, AC CENEVAL 2010). The EGEL assesses knowledge in several content areas of BA.<sup>d</sup> In Project WiwiKom, the test was translated into German and adapted according to the Test Adaptation Guidelines to ensure a high quality adaptation (American Educational Research Association, American Psychological Association and National Council on Measurement in Education AERA 2004; International Test Commission ITC 2010). The test comprises 250 items representing commonly assessed business content areas of human resources, finance, accounting, marketing, and organization and management. The items are generally in closed-ended format with various classic and complex multiple-choice formats. The EGEL was developed to cover a diverse and representative number of situational contexts that business students might encounter in their later work life. Each item comprises a situational item context and a question which refers to a situation within a company, as well as four different response options, including one attractor and three distractors. Thus, test-takers had to make the right decision within the professional context of a company. The test was

developed in a joint effort by company representatives and researchers (Centro Nacional de Evaluación para la Educación Superior, AC CENEVAL 2010). The EGEL describes typical decision situations in a professional corporate context.

In the EGEL, 28 items were originally developed to measure FK. After thorough translation, adaptation, and validation processes, 24 of the 28 items remained. Four items had to be rejected because they could not be adapted to the specific cultural and curricular context in Germany. At the same time, six new items were developed for previously under-represented content, situations and cognitive levels. After the first pilot study in 2012 ( $N = 962$ ), some items were revised and evaluated in expert interviews ( $N = 32$ ) and in an online rating ( $N = 78$ ). The experts were asked to estimate the extent to which the situational item contexts were representative of the prospective professional life of university graduates. After successful validation, all 30 items about FK were used in the main surveys.

### Data and statistical analysis

In the first major field survey, in the winter term of 2012, the 30 items were administered to bachelor level degree students from 23 universities and universities of applied sciences in Germany. In addition to FK, knowledge of human resources, accounting, marketing, and organization and management as well as of microeconomics and macroeconomics was assessed using the Test of Understanding in College Economics (TUCE; Walstad and Rebeck 2008). These further content areas were covered in 220 items. To ensure that the students could respond to a sufficient number of the financial items despite limited test-taking time, the project team used a multiple matrix design. The booklet design consisted of several complex Youden square designs (Frey et al. 2009), including 42 booklets each with three item clusters, 10 items per cluster. To confirm the item fit to the assumed levels of difficulties, each item was assigned to a level of difficulty. Therefore, nine items were assigned to the first level (understanding basic financial concepts in a company), 14 items to the second level (analyzing static concepts and static investment appraisal methods) and six items to the third level (applying and analyzing dynamic appraisal methods).

Altogether 3,873 students were assessed in this survey. The items about finance were answered by 773 students from 23 institutions of higher education (see Table 2), who formed the sample for the subsequent analyses. At the time of the survey, 24.2% of the students were in their first year of studies, while the rest were more advanced in their studies. The share of students who did not indicate their study progress was 5.4%. Approximately 20% of the students had completed commercial VET, and 13% had graduated from a commercial upper secondary school.

Overall, there were few missing values.<sup>e</sup> Missing values in the control variables (mother tongue, commercial upper secondary school attended, final school grade, finance and mathematics courses completed) were replaced using multiple imputation; five imputations were generated for each missing value.

The collected data were analyzed using models from classical test theory and item response theory (IRT). In contrast to classical test theory, IRT enables the testing of whether the sum score functions as a sufficient statistic and if it can represent a latent trait, such as FK. Furthermore, due to the booklet design, latent modeling enables

**Table 2 Sample statistics of students with or without commercial vocational and educational training**

<b>Variables</b>	<b>With commercial VET N = 155 Number (%)</b>	<b>Without commercial VET N = 618 Number (%)</b>
Gender		
Male	77 (49.68)	314 (50.81)
Female	78 (50.32)	304 (49.19)
Mother Tongue		
German	137 (88.39)	519 (84.39)
Foreign	18 (11.61)	96 (15.61)
Mathematics Course		
Completed	123 (79.35)	439 (75.43)
Not Completed	26 (16.77)	143 (24.57)
Finance Course		
Completed	82 (52.90)	302 (54.03)
Not Completed	61 (39.35)	257 (45.97)
Upper Secondary School		
Commercial	23 (14.84)	92 (16.11)
Other	124 (80.00)	479 (83.89)
Type of Institution		
University	110 (70.97)	539 (87.22)
University of Applied Sciences	45 (29.03)	79 (12.78)
<b>Variables</b>	<b>With Commercial VET N = 155 Mean (SD)</b>	<b>Without Commercial VET N = 618 Mean (SD)</b>
Final School Grade*	2.59 (.51)	2.34 (.53)
Number of Semesters	3.98 (2.04)	3.76 (2.19)

\* Final school grades range between 1 (best grade) and 4 (lowest pass grade).

comparisons between students who responded to different items. This is an advantage of IRT, since estimations of the person and item parameters are invariant to each other and can be scaled together so that the probability of a correct response to an item depends on the item difficulty and the test-taker’s ability level only (Gonzalez and Rutkowski 2010). As a reasonable alternative to the one-parameter Rasch model (1980), we could have used the two-parameter IRT model. However, due to the multiple matrix design, there could have been an insufficient number of responses per item (Eggen 2008). Under the given conditions, the Rasch model allowed a more stable estimate of the person and item parameters, and an easier interpretation of the scale (Wilson 2005). Based on the results of the Rasch model, we used multilevel modeling to identify influence factors on FK.

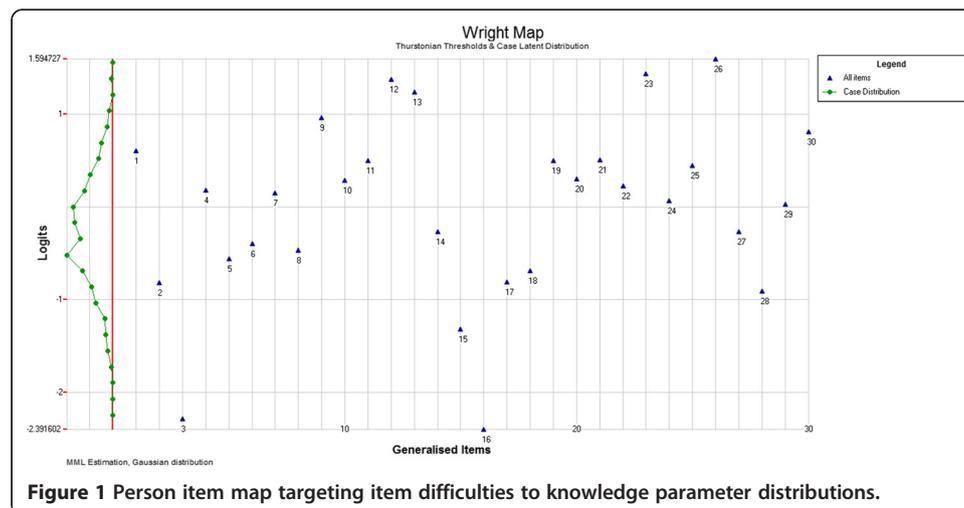
**Results and discussion**

First, the items were coded dichotomously (correct = 1 and incorrect = 0) and were entered into the multiple matrix design. Then, analyses were conducted based on item response methods. Using the ConQuest software version 3.0.1, we adjusted a one-parameter Rasch model (Adams et al. 2012) and tested for fit to the data. Item difficulty and person ability were determined using IRT based on logits, which are estimated using the marginal maximum likelihood analysis. The average item difficulty was 0.27

logits (SD = 0.96); the average person ability was  $-0.27$  logits. On average, students answered 13 of 30 items correctly (SD = 5.01). This numerical deviation indicates that the items were generally slightly too difficult for the students. Figure 1 shows the corresponding person item map, targeting the difficulty of items 1 to 30 to the person ability. Each  $x$  represents 1.3 people; both were measured in logits. Overall, the items were an adequate representation of the scale, as the pool included both very easy and very difficult items. However, there were only three items (3, 15, 16) for the lowest knowledge Level 1 ( $< -1$  logit), which means that the test provided a less accurate assessment for subjects with a very low level of knowledge. At the other end of the scale, there were more difficult items. Four items (12, 13, 23, 26) could not be matched to a corresponding ability of any of the students, though the contents of these items and the requirements were covered in the curricula of all universities. All those items required abstract mathematical knowledge, so one reason might be that abstract economic reasoning using mathematics requires a higher cognitive effort than verbally represented financial items. It is also interesting because some of the experts asked for more difficult items in the item sample. However, these items were too difficult for the sample, and further analyses are necessary to determine whether these items can be answered by more advanced, masters degree level students.

The quality of single items was examined in several item-specific analyses, including analyses of item fit statistics, item characteristic curves, and item-total correlations. Item fit statistics can be calculated using either the maximum likelihood or model-based residuals. In ConQuest, item fit statistics are indicated as weighted mean square (wMNSQ). An ideal fit of model to data is indicated by an expected value of 1, although usually a tolerance interval is defined for values between 0.8 and 1.2 (Bond and Fox 2007). Values less than 1 indicate that the data fits the model better than expected, while values above 1 indicate an overfit. For the financial items, the wMNSQ values ranged from 0.96 to 1.06. Hence, all items fit the model.

Furthermore, item fit was tested according to inferential statistics by converting the wMNSQ values in ConQuest into a standard normal distribution. The t-test indicated significant deviations from the model at a significance level of 5% for t-values outside the interval  $[-1.96, 1.96]$  (Wu and Adams 2007). Among the financial items, only item



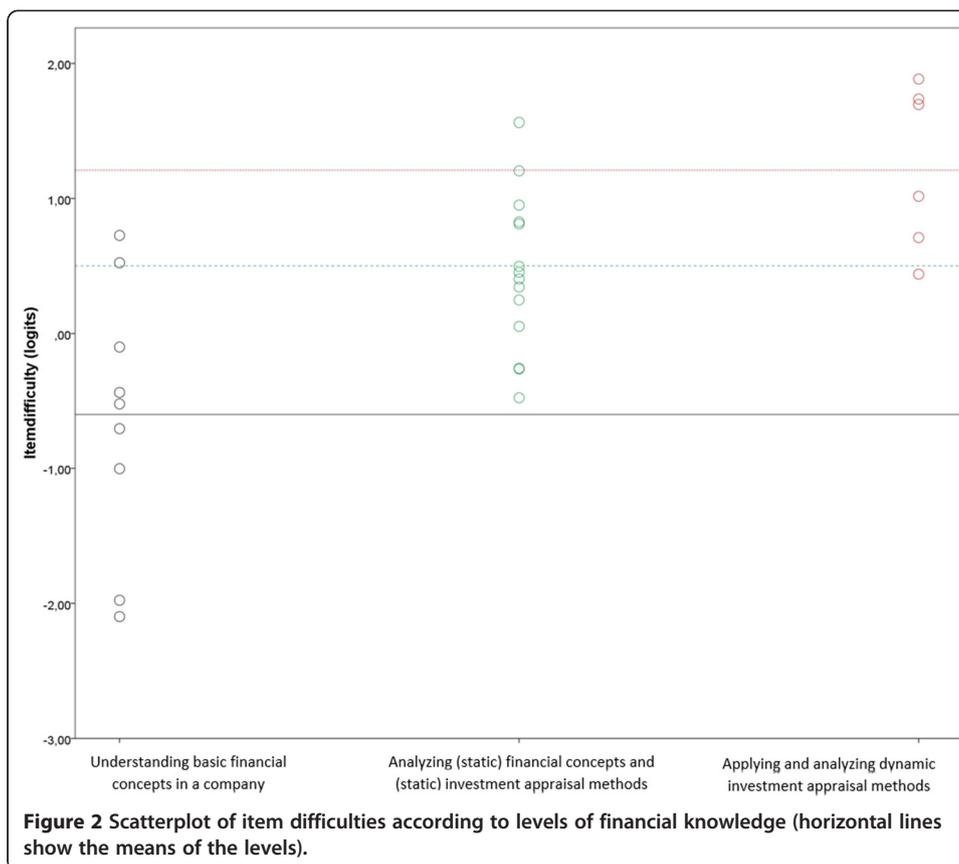
27 showed a significant deviation from the model with a  $wMNSQ = 1.06$  and  $t = 2.5$ ; therefore, it was excluded from further analyses. In addition to the specific item fit statistics of the Rasch model, item discriminatory power from classical test theory was used to estimate the fit of the items to the total score of FK. This estimation was based on the point biserial correlation between item scores and test score. The item discrimination value was expected to be positive for each of the items, since we expected, in view of the total score, that students with a higher knowledge level would respond more successfully to an item than students with a lower knowledge level. For all remaining items, positive discrimination values were confirmed with a correlation of  $> 0.30$  for 20 items and a correlation of  $> 0.20$  for nine items. The overall reliability of the 29 items was measured to be  $.658$  as the degree of internal consistency.

The Rasch model can be tested, not only with specific item fit values, but also with global fit statistics. Analyses in this regard include the likelihood-ratio test<sup>f</sup> and the Chi-squared test, which are approximately equivalent and usually give the same results. However, these tests react very sensitively and no longer follow a Chi-squared distribution when larger numbers of test items are involved, as in this case. Usually, it is necessary to use further simulation-based methods, such as bootstrapping, in order to gain evidence of the fit of the Rasch model to the available data. In bootstrapping, the initially estimated Rasch parameters are used to simulate further datasets that fit the Rasch model. Then, the Chi-squared values are calculated for the simulated datasets in order to compare them to the Chi-squared value of the originally observed data (von Davier 1997). As the ConQuest software does not offer calculations of any criteria of global fit, we used the R software with the ltm package for the likelihood-ratio test (Rizopoulos 2006). While the initial p-value of the Chi-squared test was significant ( $p < 0.01$ ), this changed after a bootstrap with 400 simulated datasets ( $p > 0.10$ ), which was clearly above the common significance level of  $p < 0.05$ . Thus, the fit of the data to the Rasch model was confirmed, not only from the item-specific perspective, but also the global perspective.

### **Research Question 1: Difficulty levels and dimensionality of financial knowledge**

After the instrument was confirmed to be compliant with the requirements for a Rasch model, the next step was to examine whether the test provided an adequate representation of the cognitive levels of FK identified in the content analyses and subject-didactic analyses. The remaining 29 items were assigned by two independent raters to one of the four levels. The mapping of items to difficulty levels was highly reliable, as the calculation of Kappa confirmed a high agreement ( $K = 0.886$ ,  $t = 6.463$ ,  $p < 0.01$ ). Eventually, nine items were assigned to Level 1 (understanding basic financial concepts in a company), 14 items were assigned to Level 2 (analyzing static concepts and static investment appraisal methods), and six items were assigned to Level 3 (applying and analyzing dynamic investment appraisal methods). Level 4 could not be included in the analysis, so further items need to be created in future research. In Figure 2, the items for Levels 1, 2, and 3 are plotted according to their item difficulty.

With regard to item difficulty, the three levels are not entirely separable, and so overlap to a certain degree. This is not surprising since item difficulty is not contingent solely on the modeled cognitive complexity, but also on numerous other variables, such



as item format and share of numerical values in an item. Figure 2 shows clearly how well the increasing item difficulty represents the modeled hierarchy in the difficulty levels. The average item difficulty was  $-0.62$  logits for Level 1,  $0.45$  logits for Level 2, and  $1.25$  logits for Level 3. The discriminatory power of the different levels in the model is also shown in Figure 2 from the fact that none of the items from Levels 2 and 3 had a difficulty below the arithmetic mean of Level 1. There was also only one item from Level 2 that had a difficulty above the mean of Level 3, and none of the items from Level 3 had a difficulty below the mean of Level 2. In summary, the first three theoretically based levels of difficulty are, on average, reflected by the assigned items. Therefore, the values range of item difficulties and the individual steps overlap each other.

Empirically, we confirmed the differences through inferential statistics by comparison in a *t*-test. The differences between the means of all three levels were significant with large effect sizes, measured with Cohen’s *d*. The values were as follows for the difference in the means of Levels 1 and 2:  $t = -3.326$ ,  $df = 21$ ,  $p < 0.01$ ,  $d = 1.42$ . For the difference between the means of Levels 2 and 3, the values were as follows:  $t = -2.756$ ,  $df = 18$ ,  $p < 0.05$ ,  $d = 1.34$ . For the difference between Levels 1 and 3, the values were as follows:  $t = -4.157$ ,  $df = 13$ ,  $p < 0.01$ ,  $d = 2.19$ .<sup>8</sup> The following analyses were conducted to verify empirically the dimensional structure of FK with regard to content. Our hypothesis was that all levels referred to the financial decision process and therefore represented one dimension. In contrast, if the results indicated different latent subdimensions, we would need to use more than one scale to prevent compromised results in the parameter estimations. We examined the dimensional structure of FK by testing

a one-dimensional model in comparison to a three-dimensional model with the observed levels included as subdimensions. The models were compared based on the Akaike Information Criterion (AIC; Akaike 1973), the Bayesian Information Criterion (BIC; Schwarz 1978), and the Consistent Akaike Information Criterion (CAIC; Bozdogan 1987). The values of these criteria must be interpreted not as absolute values, but in relation to the comparison model. For all three criteria, a smaller value always indicates a better fit of the model to the data. The comparison of the information parameters (see Table 3) shows that the values of the AIC, the BIC and the CAIC were smaller for the one-dimensional model. Thus, the one-dimensional model provided a better fit to the parameters to be estimated than the three-dimensional model.

In addition to these tests, correlations between the different levels can be analyzed as well, to answer the question of whether a one-dimensional or a multi-dimensional model provides a better fit. In a one-dimensional model, the scores of the latent variable on each level are expected to correlate clearly. Students who have mastered Level 2b and are able to make reasonable decisions at this level should also have high scores on the latent variable at the lower levels. Students who can hardly identify value streams in a company at Level 1, that is, those who are barely able to identify financial information, should have low scores at the higher cognitive process levels as well. Thus, a high correlation between the subjects' scores on the latent variable would indicate that there is only one latent variable with different levels of difficulty. In the model, the latent correlations were 0.83 between Levels 1 and 2a; 0.76 between Levels 1 and 2b; and 0.81 between Levels 2a and 2b. Both the criteria for the model comparison and the correlations between the levels indicated a model with one single latent scale with different levels, which also was in line with the theoretical considerations as to how the content difficulty increased. In summary, the increasing difficulty in financial decision-making in a company must be regarded as one content dimension of a construct with different levels rather than separate content subdimensions. This supports the theoretical assumptions. Accordingly, first, financial information must be identified in terms of value streams. Only when this has been accomplished to a sufficient degree can this information be analyzed further for evaluating financial challenges, and only after that, can decisions in complex situations be made in a well-conceived way based on FK. In the context of a company, this means that the value streams are identified, then condensed into financial indices of profitability, cost of capital, or cash flow, and then analyzed to provide a basis for decision-making. In addition to this, decision-makers must be able to apply different methods to make reasonable investment decisions under uncertainty and in view of numerous alternatives.

**Table 3 Information criteria for the one-dimensional and three-dimensional model**

	One dimension	Three dimensions
Final Deviance (-2log likelihood)	12900.16	12892.39
Number of parameters	30	35
AIC	12960.16	12962.39
BIC	13104.27	13130.51
CAIC	13134.27	13165.51

## Research Question 2: The effect of commercial vocational education and training on financial knowledge

After the theoretical structural model of FK was largely empirically confirmed, we investigated the second research question about the effect of commercial VET on the FK of bachelor level degree students, while controlling for personal and study-related influence factors. In this regard, we also analyzed whether different institutions and different types of institutions of higher education had an effect on student FK. Since we made recourse directly to the students' knowledge and since there was error variance in the estimation due to the booklet design, we estimated five plausible values of the students' latent knowledge score (Mislevy 1991). Therefore we used the five imputed datasets with the complete individual variables and computed a multiple regression using both the item scores and the imputed independent variables for the estimation (Rubin 1987). Hence, the following calculations were based on 25 datasets for each student, resulting from the five multiple imputations of the control data multiplied by the subsequently estimated five plausible values for each imputation. In the following, we report only the pooled results from the calculations with the 25 datasets.

Given the structure of our data, consisting of students nested in institutions of higher education, we used multilevel analysis to examine its influence on FK. Analyses of both the raw data and the imputed data indicated that the interclass correlation was quite small with approximately 6.8% (see Model 1 in Table 4). Thus, the level of FK differed only slightly between institutions of higher education. Hence, the variance should be explained mainly by personal influence factors. In the subsequent models, we added influence factors at the personal and institutional level. Our focus was on the effect of prior commercial VET, which we added first before including further personal and study-related influence factors.

In Model 2, we examined the influence of commercial VET alone on the knowledge score. The comparison shows that students who had obtained commercial VET prior to their studies scored higher by 0.276 logits than students without commercial VET. In Model 3, we added the personal factors, such as gender, mother tongue, commercial upper secondary school attended, and final school grade. Model 3 did not include the influence factors related to the studies in higher education. The analysis showed that the effect of commercial VET increased slightly ( $\beta = 0.327$ ), when the other personal factors were controlled. Furthermore, attending a commercially specialized upper secondary school and a better final school grade had significant positive effects on FK.

In Model 4, we also added study-related variables, such as completion of a course on finance or mathematics, and study progress in number of semesters. As expected, students who had completed a finance or mathematics course scored significantly higher. When we controlled for completion of these courses, the number of semesters had no effect on FK.<sup>h</sup> Even when completion of these learning opportunities in higher education was controlled, students with prior commercial VET still scored significantly better, by 0.329 logits, which illustrates the great relevance of VET for the acquisition of FK in higher education.

In Model 5, we added a variable to the institutional level that differentiated universities from universities of applied sciences. The analysis showed that students from universities scored slightly, but significantly, higher than students from universities of applied sciences. This variable explained almost the entire variance at the institutional

**Table 4 Multilevel analysis of influence factors on financial knowledge**

Parameter	Model 1	Model 2	Model 3	Model 4	Model 5
Fixed effects					
Intercept	-.244*** (.045)	-.309*** (.045)	-.485*** (.053)	-.789*** (.072)	-.821*** (.070)
Commercial VET (1 = yes)		.276*** (.060)	.327*** (.060)	.329*** (.058)	.315*** (.059)
Gender (1 = male)			.299*** (.050)	.299*** (.049)	.294*** (.049)
Mother Tongue (1 = not German)			-.105 (.063)	-.114 (.060)	-.118* (.060)
Final School Grade			-.142** (.044)	-.165*** (.043)	-.166*** (.042)
Commercial Upper Secondary School Attended			.200*** (.060)	.188*** (.058)	.186*** (.058)
Finance Course Completed				.187*** (.056)	.207*** (.055)
Mathematics Course Completed				.187** (.064)	.176** (.063)
Number of Semesters				.021 (.014)	.017 (.014)
Type of Institution of Higher Education (1 = University)					.138* (.061)
Var(Intercept), $u_{0j}$	0.022*	0.017*	0.016*	0.000	0.000
Var(Level 1), $r_{ij}$	0.295*	0.284*	0.251*	0.230*	0.229*
Pseudo-R <sup>2</sup> (Institutional)		0.223	0.280	0.993	0.999
Pseudo-R <sup>2</sup> (Personal)		0.037	0.150	0.220	0.222

Note. \*p ≤ .05 \*\*p ≤ .01 \*\*\*p ≤ .001. Values in brackets indicate standard error.

level. Thus, in our sample, only the type of institution was relevant for student FK, not the specific university or university of applied sciences they attended.

The variables at the personal level explained approximately 23% of the variance in Model 5 and approximately 100% at the contextual level. This can be explained by three factors. First, the variance at the context level is relatively low, thus the addition of few covariates, which are potentially declarative, cause a high variance explanation. Second, contextual covariates were included, which can readily explain the differences in test scores between the universities. Third covariates were included to show, not only the variation within one university but also between various universities. Overall, the variables in Model 5 explained approximately 27.5% of variance of the knowledge score.

**Conclusion**

The theoretical model of FK was confirmed empirically for three hierarchical levels. Accordingly, students first learn basic financial concepts before they can calculate and

analyze figures based on static methods. This, in turn, is the prerequisite for calculating figures based on dynamic methods, taking into account risk, and making professional and long-term financial decisions in a company. The adapted and further developed test instrument is suitable for representing the knowledge levels. Therefore, the effects of time, insecurity and level of abstraction of the difficulty of the task need to be investigated in further studies. A possible next step would be another construct validation which is based on predictions about the item difficulties through the assumed criteria (Hartig and Frey 2012). However, it is still limited in the sense that the items do not represent all four levels of the theoretical model of FK. The instrument does not yet represent Level 4 (creating finance plans), the highest level in the model, which should be the focus of further studies. It is important to note that this level also connects to other business content areas, which increases considerably the difficulty of decision-making and related items.

There was an important influence of commercial VET on student FK, which was in line with expectations, considering that, according to the test definition, the test is supposed to assess financial decision-making in corporate situations. Our results suggest that students who have obtained commercial VET may draw on their practical experience when responding to the test items. It might be easier for them to project themselves into the corporate context, which would enable them to respond more successfully to the items even if they had acquired the same knowledge in higher education as their peers. We consider this finding an indicator of sufficient practical relevance to the test instrument. Moreover, this effect illustrates a special quality of the dual commercial VET in Germany, as students apparently acquire knowledge relevant for decision-making during commercial VET, and the resulting edge cannot be compensated for by learning opportunities during bachelor level degree studies.

Interestingly, there was a significant difference in the share of students with commercial VET between universities and universities of applied sciences. The share of students who had completed commercial VET was only 16.95% at universities but 36.29% at universities of applied sciences ( $\chi^2 = 24.294$ ,  $p < 0.001$ ).<sup>1</sup>

We assume that the influence of commercial VET on student knowledge is still somewhat underestimated. In our study, it was one of the most influential predictors of the level of FK, and therefore, its influence and functioning should be analyzed in greater detail in further studies. A relevant question would be how the acquisition of FK in tertiary education is influenced by different parts of commercial VET, for example, by practical training in a company compared to school-based education.

## Endnotes

<sup>a</sup> Studies conducted during the financial crisis show that even people with (prior) knowledge of BA often have an incorrect or superficial understanding of the causes of the financial crisis (Leiser et al. 2010).

<sup>b</sup> The difference between these two common types of institutions of higher education in Germany is that universities aim mainly to provide academic education while universities of applied sciences are more practically oriented (e.g., Nickel 2011).

<sup>c</sup> Project WiwiKom is the acronym of the research 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 (EGEL/TUCE)’. The project is funded by the German Federal Ministry of Education and Research. For more information, see <http://www.wiwi-kompetenz.de/eng>.

<sup>d</sup> On the state of research on test instruments for assessing professional knowledge of BA, see Zlatkin-Troitschanskaia et al. (2014).

<sup>e</sup> Mother tongue (three missing, 0.39%); mathematics course completed (42 missing, 5.43%); finance course completed (71 missing, 9.18%); commercial upper secondary school attended (55 missing, 7.12%); final school grade (58 missing, 7.50%); number of semesters (16 missing, 2.07%).

<sup>f</sup> In the likelihood-ratio test, a hierarchically subordinate model (numerator) is compared to a hierarchically superordinate model (denominator).

<sup>g</sup> For Cohen’s *d* (Cohen 1988), effect sizes above 0.9 usually are considered large. In this case, all values were above 0.9.

<sup>h</sup> When completion of these courses was excluded from the model, the number of semesters had a significant effect. Thus, the factors of courses completed and number of semesters provided similar information. The mean correlation was 0.56, indicating that there was no multicollinearity, which could have compromised the regression results.

<sup>i</sup> There was also a difference in the share of students who had attended a commercial upper secondary school between universities (15.14%) and universities of applied sciences (20.51%), but it was not significant ( $\chi^2 = 2.101$ ,  $p = 0.147$ ). Due to data collections from particular courses at several universities, some subsamples show an averagely higher study progress (semester of study and attended courses) than other ones.

#### Abbreviations

AIC: Akaike information criterion; BA: Business administration; BAKT: Business administration knowledge test; BIC: Bayesian information criterion; CAIC: Consistent akaike information criterion; EGEL: Examen General para el Egreso de la Licenciatura Administración; FK: Financial knowledge; ILLEV: Innovative teach-study network in academic higher education; IRT: Item response theory; PC: Professional competence; VET: Vocational education and training; wMNSQ: Weighted mean square.

#### Competing interests

The authors declare that they have no competing interests.

#### Authors’ contributions

All authors contributed substantially to this work. OZT and MF conceptualized and raised funding for the WiwiKom project. OZT developed the general theoretical competence model, while MF and SB conceptualized the subject-specific levels in the different dimensions, respectively in finance in this paper. MF planned and coordinated the test adaptation and validation studies. SB planned and conducted the cognitive interviews. Data analysis for this paper was carried out by SB in consultation with MF. All authors discussed together the manuscript at all stages. All authors read and approved the final manuscript.

#### Authors’ information

Manuel Förster (MF) is an Assistant Professor at the Chair of Business Education of Johannes Gutenberg University Mainz, Germany. His research interests lie in quantitative methods, implementation of innovation and reforms in the educational system, and competence assessment in higher education for international comparisons. Manuel Förster is one of the project directors and international coordinator of the WiwiKom research project, which focuses on the modelling and measuring of competences and knowledge in business and economics in higher education. Sebastian Brückner (SB) is a research associate at the Chair of Business Education of Johannes Gutenberg University Mainz, Germany. He has been working in the WiwiKom research project and the German national research program on Modeling and Measuring Competencies in Higher Education (KoKoHs). His research focus is on cognitive diagnostic assessment and international comparative research in higher education. Professor Olga Zlatkin-Troitschanskaia (OZT) has been Chair of Business Education at the Johannes Gutenberg University Mainz, Germany, since 2006. She has directed numerous national and international externally funded

research projects, such as WiwiKom, and coordinates the national research program on Modeling and Measuring Competencies in Higher Education (KoKoHs) in Germany. Professor Olga Zlatkin-Troitschanskaia is a member of many national and international advisory and editorial boards and has served as an expert consultant to the German and Swiss national research foundations and the German and Swiss ministries of education and research.

### Acknowledgments

In addition to the project directors stated above, project directors include Wolfgang Härdle, Silvia Hansen-Schirra, and Sascha Hoffmann. In the curricular validation of the test instrument and the assessment surveys, the WiwiKom team was supported by Oliver Lauterbach, Hilde Schaeper, and Florian Aschinger.

Received: 6 August 2014 Accepted: 5 February 2015

Published online: 25 March 2015

### References

- Adams RJ, Wu ML, Wilson MR (2012) ACER ConQuest 3.0.1. ACER, Camberwell
- Akaike H (1973) Information theory and an extension of the maximum likelihood principle. In: Petrov BN, Csaki F (eds) Second international symposium on information theory. *Academiai Kiado, Budapest*, pp 267–281
- Alexander PA, Kulikowich JM, Schulze SK (1994) How subject-matter knowledge affects recall and interest. *Am Educ Res J* 31:313–337
- American Educational Research Association, American Psychological Association and National Council on Measurement in Education (AERA) (2004) Standards for educational and psychological testing. American Educational Research Association, Washington DC
- Anderson LW, Krathwohl DR (2001) A taxonomy for learning, teaching, and assessing: a revision of Bloom's taxonomy of educational objectives. Longman, New York
- Becker HP (2012) *Investition und Finanzierung [Investment and finance]*, 5th edn. Gabler, Wiesbaden
- Bond TG, Fox CM (2007) *Applying the Rasch model: Fundamental measurement in the human sciences*, 2nd edn. Lawrence Erlbaum, Mahwah NJ
- Bozdogan H (1987) Model selection and Akaike's Information Criterion (AIC): The general theory and its analytical extensions. *Psychometrika* 52(3):345–370
- Brückner S, Förster M, Zlatkin-Troitschanskaia O, Walstad W (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*. <http://www.tandfonline.com/doi/full/10.1080/03075079.2015.1004235#abstract>
- Centro Nacional de Evaluación para la Educación Superior, AC (CENEVAL) (2010) *Guía para el sustentante. examen general para el egreso de la licenciatura en administración (EGEL-ADMON) NUEVA GENERACIÓN*. CENEVAL, México
- Cohen J (1988) *Statistical power analysis for the behavioral sciences*, 2nd edn. Erlbaum, Hillsdale NJ
- Eggen TJHM (2008) Adaptive testing and item banking. In: Hartig J, Klieme E, Leutner D (eds) *Assessment of competencies in educational contexts*. Hogrefe, Göttingen, pp 215–234
- Frey A, Hartig J, Rupp AA (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
- Gonzalez E, Rutkowski L (2010) Principles of multiple matrix booklet designs and parameter recovery in large-scale assessments. In: Hastedt D, von Davier M (eds) *IERI Monograph Series: Issues and Methodologies in Large-Scale Assessments*, vol 3. IEA-ETS Research Institute, Hamburg, pp 125–156
- Größler A, Wilhelm O, Wittmann WW, Milling PM (2002) *Measuring business knowledge for personnel selection in small and medium sized companies: Abschlussbericht zum Projekt: Die Erfassung von Wirtschaftswissen zur Personalauswahl in KMU [Final report of the project Measuring economic knowledge for personnel selection in SMEs]*. University of Mannheim, Institute of SMEs and Entrepreneurship, Mannheim
- Happ R, Zlatkin-Troitschanskaia O, Förster M, Preuß D, Kuhn C, Schmidt S (2013) The research project ILLEV (Innovative Teach-Study-Network in Academic Higher Education) - a short summary of the main results. *Johannes Gutenberg University Mainz: Arbeitspapiere Wirtschaftspädagogik [working paper: business education]*. Available via DIALOG. [http://www.wipaed.uni-mainz.de/Is/ArbeitspapiereWP/gr\\_Nr\\_65.pdf](http://www.wipaed.uni-mainz.de/Is/ArbeitspapiereWP/gr_Nr_65.pdf)
- Hartig J, Frey A (2012) Konstruktvalidierung und Skalenbeschreibung in der Kompetenzdiagnostik durch die Vorhersage von Aufgabenschwierigkeiten [Using item difficulty modeling for construct validation and scale descriptions in competency diagnostics]. *Psychol Rundsch* 63(1):43–49
- International Test Commission (ITC) (2005) *International Test Commission guidelines for translating and adapting tests*. Available via DIALOG. <http://www.intestcom.org/upload/sitefiles/40.pdf>. Accessed 11 Nov 2013
- Jahn H (2007) Bachelor/Bakkalaureus. In: Brettschneider F, Wildt J (eds) *Handbuch Akkreditierung von Studiengängen – Eine Einführung für Hochschule, Politik und Berufspraxis [Handbook of accreditation of degree courses – An introduction for higher education institutions, policy makers, and professional practitioners]*, 2nd edn. Bertelsmann, Bielefeld, pp 203–212
- Krücken G (2004) Wettbewerb als Reformpolitik: De-oder Re-Regulierung des deutschen Hochschulsystems? [Competition as policy of reform: Deregulation or reregulation of the German higher education system?]. In: Beyer J, Stykow P (eds) *Gesellschaft mit beschränkter Hoffnung. Die ungewisse Aussichtslosigkeit rationaler Politik [Hope for Society Limited. Uncertain lack of prospects in rational policy making]*. VS Verlag für Sozialwissenschaften, Wiesbaden, pp 337–356
- Kuhn C, Happ R, Zlatkin-Troitschanskaia O, Beck K, Förster M, Preuß D (2014) Kompetenzentwicklung angehender Lehrkräfte im kaufmännisch-verwaltenden Bereich – Erfassung und Zusammenhänge von Fachwissen und fachdidaktischem Wissen [Competence development among future teachers in the field of business administration – assessment and relationship of content knowledge and pedagogical content knowledge]. In: Winther J, Prenzel M (eds) *Perspektiven der empirischen Berufsbildungsforschung: Kompetenz und Professionalisierung [Perspectives to empirical vocational education research: Competence and professionalization]*. *Zeitschrift für Erziehungswissenschaft [Special Issue 22]* 17(1): 149–167. doi:10.1007/s11618-013-0456-3
- Lauterbach O (2013) *Curricula wirtschaftswissenschaftlicher Studiengänge an deutschen Hochschulen und ihre Bedeutung für die Kompetenzmessung [Curricula of business and economic degree courses at institutions of higher education in*

- Germany and their relevance for competency measurement]. Presentation at the third convention of the German Consortium in Statistics, Freiburg
- Leiser D, Bourgeois-Gironde S, Benita R (2010) Human foibles or systemic failure—Lay perceptions of the 2008–2009 financial crisis. *J Socio-Econ* 39(2):132–141
- Mentele S, Heinzer S, Eberle F (2013) LINCAs im Testentwicklung zur Diagnose Kaufmännischer Kompetenz [LINCAs im test developement for commercial competence diagnostics]. Presentation at the autumn meeting 2013 of the Division for Vocational Education of the German Educational Research Association, TU Chemnitz, Chemnitz, 26 Sept 2013
- Ministerial Council for Education, Early Childhood Development and Youth Affairs (MCEECDYA) (2011) National Consumer and Financial Literacy Framework. Available via DIALOG. [http://www.bea.asn.au/cms/files/cms\\_files/content/Financial\\_literacy/NationalConsumerFinancialLiteracyFramework\\_2011.pdf](http://www.bea.asn.au/cms/files/cms_files/content/Financial_literacy/NationalConsumerFinancialLiteracyFramework_2011.pdf). Accessed 29 Jul 2014
- Minnameier G (2013) The inferential construction of knowledge in the domain of business and economics. In: Beck K, Zlatkin-Troitschanskaia O (eds) Professional and VET learning: From diagnostics to learning success, vol 2E, Proceedings in vocational education and training. Sense Publ, Rotterdam, pp 141–156
- Mislevy RJ (1991) Randomization-based inference about latent variables from complex samples. *Psychometrika* 56:177–196
- Nickel S (2011) Zwischen Kritik und Empirie – Wie wirksam ist der Bologna-Prozess? [Between critique and empirical results – How effective is the Bologna Process?]. In: Nickel S (ed) Der Bologna-Prozess aus Sicht der Hochschulforschung: Analysen und Impulse für die Praxis [The Bologna process as seen by higher education research: Analyses and incentives for practice]. CHE non-profit Centre for Higher Education, Gütersloh, pp 8–17
- Porter ME (2000) Wettbewerbsvorteile: Spitzenleistungen erreichen und behaupten [Competitive edges: achieving and maintaining top performance], 6th edn. Campus, Frankfurt
- Prätsch J, Schikorra J, Ludwig E (2003) Finanzmanagement [Financial management], 2nd edn. Hanser, Munich
- Preiss P (2005) Förderung kaufmännischer Kompetenzen mit Hilfe des wirtschaftsinstrumentellen Rechnungswesens - aus fachlich-curricularer Perspektive [Promotion of commercial competence by means of accounting – A professional curricula perspective]. In: Sembill D, Seifried J (eds) Wirtschaftswissenschaft. Rechnungswesenunterricht am Scheideweg. Lehren, lernen und prüfen [Business studies. Accounting teaching at a crossroads. Teaching, learning, and examination]. DUV Deutscher Universitätsverlag, Wiesbaden
- Rasch G (1980) Probabilistic models for some intelligence and attainment tests. The University of Chicago Press, Chicago
- Rizopoulos D (2006) Irm: An R package for latent variable modeling and item response theory analyses. *J Stat Softw* 17(5):1–25
- Rubin DB (1987) Multiple imputation for nonresponse in surveys. Wiley Classics Library, Wiley-Interscience, Hoboken NJ
- Schwarz G (1978) Estimating the dimension of a model. *Ann Stat* 6(2):461–464
- Tremblay K, Lalancette D, Roseveare D (2012) Assessment of higher education learning outcomes. Feasibility study report. Volume 1 – Design and Implementation. OECD. Available via DIALOG. <http://www.oecd.org/edu/skills-beyond-school/AHELOFSReportVolume1.pdf>. Accessed 29 Jul 2014
- von Davier M (1997) Bootstrapping goodness-of-fit statistics for sparse categorical data: Results of a Monte Carlo study. *Methods of Psychological Res* 2(2):29–48
- Walstad WB, Rebeck K (2008) The test of understanding of college economics. *Am Econ Rev* 98(2):547–551
- Walstad WB, Watts M, Rebeck K (2007) Test of understanding in college economics: Examiner's manual, 4th edn. National Council on Economic Education, New York NY
- Wilson M (2005) Constructing measures: an item response modeling approach. Erlbaum, Mahwah NJ
- Wu M, Adams R (2007) Applying the Rasch model to psycho-social measurement: A practical approach. Educational Measurement Solutions, Melbourne
- Zantow R, Dinauer J (2011) Finanzwirtschaft des Unternehmens: Die Grundlagen des modernen Finanzmanagements [Corporate finance: Basics of modern financial management], 3rd edn. Pearson, Munich
- Zlatkin-Troitschanskaia O, Förster M, Brückner S, Happ R (2014) Insights from a German assessment of business and economics competence. In: Coates H (ed) Higher education learning outcomes assessment – international perspectives. Peter Lang, Frankfurt/ Main, pp 175–197
- Zlatkin-Troitschanskaia O, Förster M, Schmidt S, Brückner S, Beck K (2015) Erwerb wirtschaftswissenschaftlicher Fachkompetenz im Studium – Eine mehrrebenenanalytische Betrachtung von hochschulischen und individuellen Einflussfaktoren [Acquisition of business and economics competence during studies in higher education – A multilevel analysis of institutional and individual influence factors]. In: Blömeke S, Zlatkin-Troitschanskaia O (eds) Kompetenzen von Studierenden [Competencies of higher education students], *Zeitschrift für Pädagogik*. 61(Beiheft 61):116–134

**Submit your manuscript to a SpringerOpen® journal and benefit from:**

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

---

Submit your next manuscript at ► [springeropen.com](http://springeropen.com)

---