Adapting and Validating the Test of Economic Literacy to Assess the Prior Economic Knowledge of First-Year Students in Business and Economic Studies in Germany

1. Relevance

Higher education systems around the world currently are moving towards greater internationalization and harmonization (Coates, 2014), for example, the bachelor/master study model is being adopted throughout Europe (for the Bologna reform, see Adelman, 2009). Consequently, there is an increase in the mobility of university students within and across countries, especially in the domain of business and economics (Brückner, Förster, Zlatkin-Troitschanskaia & Walstad, in press). This harmonization process should lead to a greater degree of compatibility among university education systems, and the curricula of universities should become more comparable. Current international comparison studies investigate whether the learning outcomes in higher education in different countries also can be compared. The assessment of higher education learning outcomes is changing in many countries due to increased focus in research on finding comparable measurement methods which can be applied on an international scale (Coates, 2014).

The aim of this paper is to develop a model of previous economic knowledge of first-year students in economics degree courses at universities in Germany using an adapted American test to determine whether the knowledge of the students in the two countries is comparable. Cross-national comparative analyses in higher education have at least two prerequisites (OECD, 2011): First, the content to be compared must be the same in the countries under investigation and valid in terms of curricula. This requires not only translation of the American Test of Economic Literacy (TEL), but also complex adaptation of the test instrument (Coyne, 2000; Hambleton, 2001) to ensure comparability of the construct across nations despite, for example, cultural differences. Second, a clear statement is needed on what previous knowledge first-year university students should have when they begin their studies. To determine what influences economic education in each country under investigation, assessments must be made with comparable test instruments that are valid in terms of curriculum in
those countries. In this paper we use the fourth edition of the TEL (TEL4; Walstad, Rebeck & Butters, 2013) and we describe the adaptation and validation process of this measurement instrument. Building on the results, we investigate the extent to which previous economic knowledge of first-year university students in Germany can be assessed in a valid way using the adapted TEL4.

2. Theoretical Criteria for Adapting and Validating Psychological Tests

Adapting a psychological test is a complex task (e.g., Zlatkin-Troitschanskaia, Förster, Brückner & Happ, 2014). Although Test Adaptation Guidelines (TAG) have been issued by the International Test Commission to ensure high quality test adaptations (Coyne, 2000; Hambleton, 2001), they provide only a rough orientation and need to be specified with regard to the content of the respective project. In our study, the TAG were used as a methodological basis for the translation and adaptation process of the TEL4.¹ The TAG are divided into four categories: (1) context, designed to ensure that assessment, of knowledge of different populations target the same aspects of theoretical constructs as well as to minimize cultural influences irrelevant to the assessment; (2) test development and adaptation, focusing on issues related to translation, data collection, and statistical analyses; (3) administration, dealing with testing procedures and issues arising when subjects have different linguistic and cultural backgrounds; and (4) documentation/score interpretations, emphasizing the importance of documenting the test and changes made during the adaptation process to ensure validity and avoid diagnostic misinterpretation.

In our study we followed the TAG and adhered to the American Educational Research Association (AERA) Standards to evaluate the validity of the TEL4. According to the AERA, standards must be met in the following five areas: test content, response processes, internal structure, influence of other variables, and consequences of testing (AERA, 2004, pp. 11-17).

1) Test content

Test content is analyzed to determine how accurately it represents theoretical constructs (AERA, 2004, p. 11). The accuracy of representation can be determined through logical or empirical analysis (AERA, 2004, p. 11). One option for an empirical analysis is to ask experts to evaluate how the test content relates to content of the respective field of study. Thus, the content of the TEL4 can be evaluated by experts from the respective subject areas of economics. This analysis is particularly important when a test is administered in a new educational context where the education system and

¹ The TAG already have been broadly used in test adaptation practice. They have been the basis for the adaptations, for example, of the third Trends in International Mathematics and Science Study (TIMSS) and the Program for International Student Assessment (PISA) (Grisay, 2003; Hambleton, 2001); however, the guidelines were specified in quite different ways in these studies.
the curriculum might differ from the original target context. Curricular differences must be taken into account when adapting and validating a test to ensure valid assessment and comparison of test outcomes among countries. In our study, we ensured that the construct of previous economic knowledge would be conceptualized and understood in a similar way in America and in Germany (AERA, 2004, p. 12). This criterion of the AERA Standards corresponds with the criterion of context in the TAG.

2) Response processes

Response processes of the test subjects are analyzed to determine whether there is a good fit between theoretical constructs and the observed item response processes. Individual response strategies are examined, and items can be revised if they are repeatedly misunderstood. When validating the TEL4, we ensured that item responses were indeed based on the intended cognitive processes and test solving strategies and that, in contrast, undesired test-taking and guessing strategies did not result in correct responses (AERA, 2004, pp. 12-13).

3) Internal structure

The internal structure of a test is analyzed to determine the extent to which a construct is coherently represented by the relations between single items or various parts of a test. How the internal structure is analyzed and how the results are interpreted depends on the aim of the test, that is, on the initially assumed structure. For instance, a one-dimensional construct or test is expected to have fairly homogenous items (AERA, 2004, p.13). We assumed the content of the TEL4 comprised basic economic concepts for first-year students of economics. A further assumption was that, as a rule, regular schools and vocational schools often do not differentiate between microeconomics and macroeconomics. Economics often is taught unsystematically and in other subjects in secondary school, and the difference between microeconomics and macroeconomics usually is first introduced in higher education. We assume, therefore, that the content being taught is not categorized or presented in a multifaceted manner, but rather is seen as a one-dimensional construct of economic knowledge.

4) Influence of other variables

Another validity criterion lies in the relationship of the test results to external variables, which must be analyzed according to the elements of a nomological network (AERA, 2004, pp. 13-16; Cronbach & Meehl, 1955). External variables may be personal or group-related, and may reveal a convergent or discriminant relationship between the construct and the respective variable. For instance, in line with

---

2 In most secondary schools in Germany economics is not a subject offered. Economic concepts often are taught in social science.
the expert-novice paradigm (Ericsson, 2008), we assumed that students who obtained vocational training before they started their studies would score higher on the economic knowledge test than students who did not.

5) Consequences of testing

The consequences of testing refer to the conclusions drawn from the test scores, which allow evaluation of individual students or groups of students. For the TEL4, we needed to document sufficiently the findings from the test development process and provide a manual that would explain the possible uses of the test, for example in higher education economics, so that those administering the test would not draw inappropriate conclusions or administer the test in an inappropriate way (AERA, 2004, pp. 16-17).

3. Adaptation and Validation Process of the TEL4 and Preliminary Results

In the following, we present the adaptation and validation process and highlight the results indicating that the TEL4 is a valid tool for assessing the previous economic content knowledge of students in higher education in Germany. We have classified the validity analyses according to the above-mentioned validity criteria although some of the analyses can be connected to more than one validity criterion.

3.1 Translation of the TEL4 and the Adaptation Process

First, the TEL4 was translated into German by professional translators specialized in economics to ensure the translation was of very high quality. There are two parallel versions (A & B) of the TEL4 (Walstad et al., 2013). Both are multiple-choice tests containing 45 questions each with four answers from which to choose. The questions refer to one short description of an economic situation. The test takers are informed beforehand that there is only one possible answer for each question. To compare the two TEL4 versions (Walstad et al., 2013), there are 10 identical anchor items on each.

During the translation and adaptation process, the researchers were able to draw on their experience with the adaptation of the second version of the TEL (Soper & Walstad, 1987), which already had been adapted to meet German requirements by Beck and Krumm (1991) and had been applied in the ILLEV Project (Zlatkin-Troitschanskaia, Förster & Kuhn, 2013; Happ & Zlatkin-Troitschanskaia, 2014). The same was the case for the Test of Understanding College Economics
(TUCE; Walstad, Watts & Rebeck, 2007), which had been adapted from American English to German and applied in the WiwiKom\(^3\) project (Zlatkin-Troitschanskaia, Förster, Brückner & Happ, 2014).

The method employed by Beck and Krumm (1991) in the translation and adaptation process involved translating word by word and sentence by sentence, and reporting unavoidable grammatical and lexicographical problems. Some English sentence structures were almost impossible to translate word-by-word. For certain English terms, there was no exact German translation; therefore, the translation had to convey the meaning as closely as possible. The problem of equivalence, that is, how closely a translation must adhere to the original, has been discussed extensively in translation studies, for instance, by Koller (1979) and Albrecht (1990). In more recent years, translation scholars have turned away from equivalence-based translation models in favor of target-oriented approaches, such as the functionalist and action-oriented translation models proposed by Reiß and Vermeer (2013/1984). These approaches take into account the linguistic and cultural characteristics of the source and target texts and today are considered more suitable for producing high-quality translations. Hence, target-oriented approaches also were used in our study. First, the source texts were translated with annotations by a certified translation service provider specialized in the field. Linguistic supervision and quality assurance was provided by the English Department of the Faculty of Translation Studies, Linguistics, and Cultural Studies of the Johannes Gutenberg-University Mainz. The entire translation process met the highest academic quality standards. For instance, terminology management was used to optimize consistency, and proofreading was done by two professionals to assure quality. The translation was revised after several feedback sessions during which experts from the field of economics reviewed the test and reported any field-related shortcomings of the content. Experts in economics and in translation studies came together to revise the translation and validate a final version for use in the field (see D1-D3 of the TAG; Hambleton, 2001). The result was that only a few items required further cultural adaptation and all of the items were successfully adapted to a German context.

### 3.2 Validation process

**Validation of the TEL4 content**

To ensure content and curricular validity of the TEL4, a variety of analytical techniques were used. The initial database for this study included the analyses of curriculum content gathered and

---

\(^3\) WiwiKom is the abbreviation of a project conducted in Germany called ‘Modeling and measuring competencies in business and economics among students and graduates by adapting and further developing existing American and Latin American measuring instruments’. For more information, see http://www.wiwi-kompetenz.de/eng/index.php.
researched for the WiwiKom project. During the WiwiKom project, 96 degree courses were analyzed at 40 universities and 24 universities of applied sciences in Germany\(^4\), including the curricula of the largest business and economics faculties at higher education institutions. The WiwiKom team reviewed study program descriptions and module manuals to determine what constituted a core curriculum, that is, which courses were taught at most or all higher education institutions in Germany. Using those findings, we verified for our study of the TEL4 whether the content of the test items were part of the curricula of universities and universities of applied sciences in Germany. Some of the studies conducted on higher education show that the information provided by universities and universities of applied sciences in program descriptions and in module manuals often are not in line with actual teaching content. Therefore, as a second step, the content of economics curricula in Germany was investigated using textbook analyses.\(^5\) Textbooks which had direct relevance to the target groups (first-year university students of business and economics) were reviewed and, if suitable, used for the textbook analyses. For the purpose of analyses in our study, we drew on widely used textbooks with a benchmark status for our target group samples (more details in Happ, in prep.).

Our evaluation of the textbooks, which we considered to be especially important, revealed that the majority of the books had a distinctive international character. The textbooks published in American English usually had been translated into many other languages, allowing them to be distributed internationally, including in Germany. It is clear that the list of pertinent standard literature as presented by Lopus and Paringer (2012, p. 297), that is, from an American perspective, also is becoming increasingly used in Germany. In this study the text of the following four textbooks was analyzed: i) *Economics* by Samuelson and Nordhaus (2010), which is particularly valuable, and now has become a standard reference internationally; ii) *Economics - Principles, Problems and Policies* by McConnell, Brue and Flynn (2012); iii) *Principles of Economics* by Mankiw (2012), now in its 5th edition by Mankiw and Taylor (2012), which has been translated into German and is used in universities throughout Germany; and iv) *Economics* by Krugman and Wells (2010), which comes from America and has been translated into German.

For the textual analysis, the German adaptation of the content on the TEL4 was compared to that in economic textbooks designed for first-year students. It became clear with one look at the table of contents in these economic textbooks that there is considerable similarity in the essential content

---

\(^4\) In Germany, the two common types of higher education institutions are universities and universities of applied sciences. Universities mainly aim to provide academic education, while universities of applied sciences are more practically oriented.

\(^5\) The approach taken to analyzing these textbooks involved considering theory as well as the information gathered during interviews with experts. The approach taken to analyses was textual analysis (e.g., Watson’s contribution, 1997).
areas covered. It is only on a more detailed level that the textbooks begin to differentiate. Nonetheless, the differences have less to do with the content areas than with the depth to which topics are covered and the emphasis placed on the topics. In most of the textbooks analyzed there was at least one trisection of *basic principles of economics, microeconomics* and *macroeconomics*. This corresponds to the OECD’s focus of analysis (2011, p. 14), which begins with “general degrees in economics,” continues with “specialized degrees in economics” (microeconomics and macroeconomics), and ends, as a kind of advanced perspective, with “combined degrees in economics taught using an interdisciplinary approach.”

The assignments of the TEL4 are based on the 20 content standards in economics of the Council for Economic Education (CEE; CEE, 2010). In accordance with these standards, key concepts were designed that led to the different items (Walstad et al., 2013, p. 302). After reviewing the textbooks, it became clear that the basic concepts related to the items in the textbooks under investigation were in the *basic principles* section, which in turn points to a one-dimensional structuring of knowledge among first-year university students. Because of this and our findings from the textbook analysis, the assignments on TEL4 also can be considered valid in terms of curriculum for first-year university students of economics. Furthermore, findings from the analysis of program descriptions or module manuals as well as the textbooks prove that the TEL4 is valid in terms of content and curriculum for first-year university students in Germany.

**Validation of the internal structure and relations to other variables**

The German version of the TEL4 was tested empirically for the first time in a major field survey with 1,403 first-year university students. Because the test takers were beginning their first semester, it is unlikely that they had learned any economic concepts from university studies. First-year students from 10 universities in Germany participated in the survey: 65.8 % from regular universities and 34.2 % from universities of applied sciences. Female students represented 48.4 % of the test takers. While 705 participants did test version A, 698 did test version B.

We examined the internal structure of the test using this sample. The question about the dimensionality of specialized knowledge was evaluated by means of confirmatory factor analyses. The calculations were made using Mplus (Muthén & Muthén 1998-2014) and WLSMV Estimator, which made estimates for versions A and B of the test. For both TEL4 A (Estimate: WLSMV; N = 1403;)

6 In addition to this view on content, an outline of assignments from a cognitive perspective also is made with emphasis on Bloom’s taxonomy (1956). Bloom’s outline is then followed by a test assessing three levels of cognition: knowledge, comprehension and application. For each of these levels, assignments were taken from the TEL4 (Walstad, Rebeck & Butters 2013, p. 302).
A good match of the model to the data was found. Using a model with three dimensions separating *basics in economics*, *microeconomics* and *macroeconomics* could not provide a significantly better fit. This means that microeconomics and macroeconomics cannot be systematically separated based on this data. This can be explained by the assumption that the comprehending fundamental principles (e.g., shortage) is essential to understanding more advanced concepts in both microeconomics and macroeconomics. It still holds true that all items loaded significantly positively on the latent factors. All factor loadings from test version A were larger than 0.15 while only two factor loadings from test version B were smaller than 0.15. Therefore, the analyses to come will be conducted using the total score for economic knowledge.  

We also analyzed whether there were differences in the knowledge scores of the two groups and if so, whether they were due to influences of other variables, for instance, gender or the completion of commercial vocational training prior to studies. Such analyses of the relationship between test scores and external variables provided indications for the convergent or discriminant validity of the test. Our hypotheses were tested using analyses taken from classical test theory and item response theory. As expected, the TEL4 had convergent validity, as it provided a good measure of the positive effect of previous education: completing vocational training before starting university studies had a positive effect on economic knowledge. Gender also had a strong impact on the test scores: On both the original American test (see Walstad et. al., 2013) and the adapted German test, male students performed better than female students when other relevant influential factors were evaluated. This might compromise the fairness and validity of the test and require further analyses.

When the performance of the test takers of both German TEL4 versions was assessed, it became apparent that the scores achieved on version A of the test were lower than those on version B. The difference in mean value was just over one point between the two versions, constituting a difference when compared to the results recorded in America (Walstad et. al., 2013, p. 18). The American students reached almost the same mean values between the two versions. Accordingly, further analyses on an item level should be performed to determine which assignments in both test versions contribute to these differences among the test takers of the German test versions.

---

7 It still remains certain that first-year university students are still showing a one-dimensional structure of economic knowledge. Advanced students are able to differentiate between microeconomics and macroeconomics with a correlation of 0.70 (Zlatkin-Troitschanskaia, Förster, Schmidt & Brückner & Beck, in press). It still should be researched at what point and time different structures form in economic knowledge.
4. Conclusion and further steps

Overall, our results show that the TEL4 allows valid conclusions to be drawn about previous economic knowledge of first-year university students of business and economics in Germany. During the presentation, we will discuss in greater detail the statistical analyses conducted for the German survey. The next step will be to determine the requirements for international comparisons, for instance, between Germany and America. A feasibility study will need to compare in particular the empirical measurement models used in both countries. From the models it will be possible to draw conclusions about the degree of similarity between the underlying latent constructs in both countries. Data from the TEL4 in Germany and in America currently are being prepared for comparison. Subsequently, various confirmatory factor models and Rasch models can be applied to analyze the data from both countries with regard to different kinds of measurement invariance. The aim will be to conceptualize a joint measurement model which may allow us to estimate unbiased knowledge scores in both countries (see Förster et al. 2015 for procedures on an international comparison in the field of economics). From the knowledge scores recorded it will be possible to identify differences and similarities between Germany and America. Thereafter, further issues can be addressed, such as the extent to which the economic knowledge of students differs between the two countries, and whether there are structural differences between the countries.

References


Happ, R. (in prep.). Die Entwicklung des volkswirtschaftlichen Fachwissens bei Studierenden der Wirtschaftswissenschaften und der Wirtschaftspädagogik im Verlauf des Studiums – Eine vergleichende Längsschnittstudie [Development of the Economic Content Knowledge of Business Education Students and Business and Economics Students over the Course of Their Studies – A Comparative Longitudinal Study]. (Doctoral dissertation: Johannes Gutenberg-Universität Mainz)


