Roland HAPP\textsuperscript{1} & Olga ZLATKIN-TROITSCHANSKAIA (Mainz)

Assessing the Quality of the Bachelor Degree Courses: An Analysis of their Effects on Students’ Acquisition of Economic Content Knowledge

Abstract

Quality development in higher education (HE) relies on the assessment, evaluation, and optimisation of university teaching. Only a few studies have evaluated the new bachelor (BA) and master (MA) degree courses in Germany based on an objective measurement of key performance indicators, such as students’ knowledge acquisition. To address this research deficit, students’ content knowledge of business and economic (b&e) was assessed in a longitudinal study. Results of students in the BA/MA study model were compared to those of students in the former Diplom study model while many potential structural and personal influence factors were controlled. This paper presents in detail the findings and discusses their implications for quality assurance and performance assessment in higher education.

Keywords

Bologna reform, output and outcome orientation, quality development in higher education teaching, performance evaluation

Qualität der Bachelorstudiengänge auf dem Prüfstand: Analyse der Effekte auf den Erwerb des ökonomischen Fachwissens

Zusammenfassung


\textsuperscript{1} E-Mail: roland.happ@uni-mainz.de
1 Initial Situation and Research Challenge

In recent years, HE in Germany has been characterised by two particular developments. First, institutions have been subject to far-reaching structural changes in teaching and learning as a result of the Bologna reform (see NICKEL, 2011). These change processes were driven primarily by educational policy, which obliged HE institutions to adopt the BA/MA study model and to submit their degree courses to accreditation procedures (MOSCHNER, 2010; for aims of the Bologna reform, see BMBF, 2013). This reform of HE exacts more transparency of service provision and the implementation of standardisation and quality assurance processes (PAETZ et al., 2011).

Second, national and international competition among HE institutions has increased. This is evident, for example, in the constantly rising number of comparative studies. In Germany, there are regular institution rankings by the Centre for Higher Education (CHE); internationally, there are the annual Times Higher Education World University Rankings, the Shanghai Rankings (see e.g., DEHON et al., 2010), and the OECD’s endeavours in the “Assessment of Higher Education Learning Outcomes (AHELO)” feasibility study, which also takes into account the field of economics (see OECD, 2011).

These two partly associated developments have led to a change in the understanding of what quality of HE institutions means (see SCHMIDT-HERTHA, 2011). Previously, it was believed that quality was inherent to HE institutions per se and that the system assured quality by its mere self-conception and structure (LASKE et al., 2000). Today, in contrast, increasing attention is being paid to the need for systematic quality assurance and evaluation (see KRÜCKEN et al., 2010). One aim of HE institutions is to prove in a transparent way that quality objectives are being met in accordance with study programmes and curricula, which should strengthen their competitiveness (PAETZ et al., 2011). According to HELMKE et al. (2000), this development can be described as a shift in orientation from quantity, in the sense of an equal and compensating distribution of resources and opportunities, towards quality and excellence, in the sense of intended effectiveness and proven achievements of HE teaching.

A major indicator of quality of academic education is the students’ study performance or learning success (see e.g., BURTON, & RAMIST, 2001). Systematic quality development in HE relies on the key components of assessment, evaluation, and optimisation of performance in HE teaching (HOPBACH, 2012). In the literature, study performance or learning success is considered a central output criterion, which has a decisive influence on the outcome, for example, on graduate students’ success transitioning into the labour market (for measuring outcomes in a programme evaluation, see SPIEL et al., 2013).
According to § 7 of the German Framework Act for HE, “[the] purpose of teaching and study is to prepare students for a field of professional activity and to impart to them the requisite specialized knowledge, skills and methods in a way appropriate to each course [...]” This is also emphasised by BRAUN, & HANNOVER (2011), who regard HE teaching as the provision of learning opportunities for the acquisition of competencies in specific and generic domains. However, with regard to the area of HE, there is a particular lack of research approaches that allow an objective assessment of students’ skills and knowledge (KUHN, & ZLATKIN-TROITSCHANSKAIA, 2011). Currently, there are very few empirically confirmed findings on the effects of the Bologna reform in general and on the effects of the BA/MA degree courses on study performance or learning success (see ZLATKIN-TROITSCHANSKAIA et al., 2012; NICKEL, 2011) in particular. The focus of previous studies usually has been on describing the structural or organisational measures of the HE reform. They have shown e.g. that, in many places, introduction of BA/MA degree courses has resulted in little change in curricular structures and has led to a division of the former degree courses into consecutive elements (WINTER, 2011). According to this perspective, the BA/MA degree courses generally would equate to the Diplom degree courses from the curricular point of view. Based on data from the ILLEV research project, this paper presents an empirical examination of the effectiveness of the Bologna reform specifically with regard to performance-related output variables, such as individual study performance or learning success. The structural changes of the Bologna process have reached a transitional stage in which the Diplom study model still exists alongside the modularised BA/MA study model. This has created historically unique conditions for a quasi-natural experiment. In the ILLEV project, students from the new and the old study models were systematically compared to provide an initial answer to the following key question:

How do the structural changes implemented in the course of the Bologna reform affect key indicators of study performance and learning success, such as students’ economic content knowledge?  

We cannot assume any systematic effects in favour of the BA/MA study model based on findings from the curricular analyses by WINTER (2011). KLOCKE, & KRÜCKEN (2013) do not speak of a comprehensive reform endeavour at all, but rather call the process a “relabeling”.

2 The ILLEV project, short for Innovative Teach-Study Network in Academic HE, was funded by the German Federal Ministry of Education and Research under grant number 01PH08013; see also http://www.wipaed.uni-mainz.de/ls/1189_ENG.HTML.php

3 We follow an extensive theoretical model of competence within the framework of the ILLEV research project. In addition to analyzing content knowledge, we gather data on further facets of competence such as motivational orientations and epistemological beliefs (ZLATKIN-TROITSCHANSKAIA et al., 2013). Due to limited space, in this paper we outline only the results of one facet of competence: content knowledge.
3 Hypothesis: Students’ economic content knowledge does not differ significantly between the two study models.

2 Research Approach

2.1 Conceptual Background

Since the 1990s, the term quality has become a catchword in HE (RINDERMANN, 2009). In the course of the Bologna reform, HE institutions have introduced and tested a variety of quality development systems (PAETZ et al., 2011). However, an increasing number of HE institutions as well as students are questioning whether such quality assurance measures for HE teaching, such as the accreditation procedures introduced approximately 10 years ago, are actually effective in assuring quality (cf. BLOSSFELD et al., 2013).

Despite its major relevance, the concept of quality has not yet been accurately specified in the literature (KROMREY, 2006). HEID (2000) defines quality not as an observable trait of an object, but the result of an evaluation process. Accordingly, the term functions as a central concept or umbrella term for a variety of perspectives, interests, intentions, and concepts. The quality development system in HE focuses primarily on the areas of teaching, research, and management (LASKE et al., 2000). Quality in the area of teaching, which is of interest here, can be measured particularly through observed changes in students’ knowledge (RINDERMANN, 2009).

With this in mind, the focus of the ILLEV project was output-oriented criteria of study performance or learning success as indicators of quality assurance. In the project, students’ b&e content knowledge, which depends not only on personal factors but also on the structural framework conditions of a HE institution, was examined (RINDERMANN, 2009).

Personal factors of students that can influence considerably quality in study performance or learning success may include previous knowledge, motivation, and epistemological beliefs. Studies on HE didactics (e. g., SCHMIDT, & TIPPELT, 2005) have shown that experienced handling of such specific personal factors profoundly influences students’ study performance. Structural framework conditions include the respective degree course, study model, and type and number of classes attended. In this paper, focus is on the structural factor of the study model, and analysis and comparison of the influence of the BA/MA and Diplom study models on students’ content knowledge.

---

4 For a comprehensive presentation of the project’s theoretical conceptions regarding the modelling and measuring of students’ professional b&e competence, see, e. g., ZLATKIN-TROITSCHANSKAIA et al. (2013).
2.2 Methods

At the university under investigation, the shift from the Diplom degree courses to the BA/MA degree courses was initiated in the winter term of 2007/2008. Accordingly, students from both study models were surveyed from 2008 to 2011 at four measuring dates with sample sizes ranging from 800 to 1,250 students. The research hypothesis formulated in section 1 was tested for a subsample. The results of BA and Diplom students were compared at two different points of their study progress, namely in the 4th and 6th semesters (see Tab. 1).

<table>
<thead>
<tr>
<th></th>
<th>N (BA students)</th>
<th>N (Diplom students)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4th semester</td>
<td>64</td>
<td>111</td>
</tr>
<tr>
<td>6th semester</td>
<td>52</td>
<td>62</td>
</tr>
</tbody>
</table>

Tab. 1: Sample sizes

First, a curricular analysis was conducted with the study regulations and module manuals for the BA degree and Diplom courses of b&e at the university in question. The curricular analysis showed that at the time of measuring, the classes and content being taught were largely comparable for the two study models. In the survey, the students were asked which classes they had attended during their studies. At the end of the 4th semester, both groups had attended classes on the basics of b&e. At the end of the 6th semester, students in both groups had attended additional advanced and specialised classes, but on the whole, their classes were comparable.

In the first step of the data analysis, means were compared with a t-test. This test served to determine whether there were significant differences in the content knowledge between the two groups. In the second step, a multiple linear regression analysis was conducted while further potential structural and personal influence factors were controlled to see whether and to what extent the structural factor of the study model could explain the variance of the dependent variable of content.

---

5 For sociodemographic information on the students surveyed for the total sample, see ZLATKIN-TROITSCHANSKAIA et al. (2013).

6 The sample analyzed here refers to the first three measuring dates in the project, which were in the winter terms of 2008/09, 2009/10, and 2010/11. The distribution of sociodemographic information (e.g., age, gender) in both subsamples was largely consistent with the total sample.

7 To ensure comparability of the attended classes, we included questions in the survey about the classes students had attended during their studies. The results showed that especially for the two subsamples of advanced students there were differences in the course offerings in both study models (for more information, see ZLATKIN-TROITSCHANSKAIA et al., 2012).
knowledge. The analysis also showed which other factors contributed to the explained variance of content knowledge in the sample.8

The students’ b&c content knowledge was assessed using two test instruments. The first was the Wirtschaftskundliche Bildungstest (WBT) by BECK, KRUMM, & DUBS (1998), which assesses “economic knowledge and thinking”. It is an adaptation of the American Test of Economic Literacy (TEL) by SOPER, & WALSTAD (1987). The latter has been used and is still widely used on a national level and an international level. Consequently, there is substantial comparative data for this test (see e. g., for Japan, JAMAOKA et al., 2007; for the U.S. BUTTERS, & ASARTA, 2011).

The WBT focuses on economic content. It includes items from the content areas of “basics of economics”, “international relations”, and “microeconomics and macroeconomics”. The measurement properties and quality criteria of the test have been well researched and validated both for the English and the German versions (see e. g., BECK et al., 1998; SOPER, & WALSTAD, 1987). Even though the developers recommend the test be used particularly in non-academic vocational training (BECK et al., 1998), in the context of the ILLEV study, it was used with university students as well, since the items were proven to differentiate very well between low and very high levels of knowledge (see e. g., FÖRSTER et al., 2012). After the items were analysed with regard to their curricular validity for the university under investigation, 33 items were submitted to a pretest with students of b&c. After further analyses of selectivities and ceiling effect, 19 test items were deemed suitable for the target group. These items were summarized in a scale on “economic knowledge” and were used in the subsequent assessments.

The second test, which complemented the WBT, was the Business Administration Knowledge Test (BAKT) by BOTHE, WILHELM, & BECK (2006), which covers business content knowledge. This test includes items related to nine different business content areas, such as finance, management, and so on. The BAKT was developed specifically for use at university. For the ILLEV study, items were selected in line with the results of a curricular analysis and of an online rating by lecturers. The selected items belonged to the content areas of marketing, accounting, and human resources.9

Curricular validity of both tests was determined by analysing the curricula of the observed degree courses. In addition, lecturers of the respective classes were surveyed about how they judged the curricular validity of the items. The test items were selected based on the results of these analyses (cf. FÖRSTER et al., 2012). The confirmatory factor analyses showed that a one-factorial model could be as-

8 In this project, various personal factors, including previous knowledge, intelligence, etc. were assessed. Due to limited space, this paper will present analyses of only intelligence and previous knowledge, which were gained during possible vocational training completed prior to university studies.

9 In the following, only the findings from the WBT will be presented to test the hypothesis. Nonetheless, analyses of results on the BAKT support these findings as well (cf. ZLATKIN-TROITSCHANSKAIA et al., 2013).
sumed for the economic items of the WBT.\textsuperscript{10} The analysis of the item difficulties showed no ceiling or floor effects for the WBT (see HAPP, in prep.).

In addition to the test items on b&c content knowledge, the survey also included a socio-biographical section to gather data on gender, mother tongue, school leaving grade, and possible vocational training completed prior to studies. Further personal parameters were included as control variables, for example, the subjects’ intelligence, intrinsic and extrinsic motivation with regard to their chosen fields of study, as well as epistemological beliefs. The students’ intelligence was assessed using the intelligence structure test I-S-T by LIEPMANN et al. (2007) with scales on analogies and number series. Furthermore, the survey included questions on the structural framework conditions. Data also was gathered on the degree course, type of study model (BA/MA or Diplom), number of semesters, and type and number of classes attended.

\section*{3 Empirical Results}

The hypothesis was tested based on the sum scores achieved on the knowledge test. Tab. 2 shows an overview of the distribution of means in the sample. For example, the BA students in the 6\textsuperscript{th} semester, who were about to finish their BA studies, scored on average 12.46 points from a maximum of 19 points in the economic content knowledge section, with each item being worth one point. This means that, on average, the students were able to solve (only) 65.5\% of the items.

<table>
<thead>
<tr>
<th>Semester</th>
<th>Mean BA students</th>
<th>Mean Diplom students</th>
</tr>
</thead>
<tbody>
<tr>
<td>WBT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4\textsuperscript{th} Semester</td>
<td>8.84 (out of 14 points)</td>
<td>8.72 (out of 14 points)</td>
</tr>
<tr>
<td>6\textsuperscript{th} Semester</td>
<td>12.46 (out of 19 points)</td>
<td>13.26 (out of 19 points)</td>
</tr>
</tbody>
</table>

Tab. 2: Means in the economic content knowledge section

A t-test for independent samples was used to determine whether the observed means of BA and Diplom students differed significantly. This version of the t-test requires a preliminary test for homogeneity of variance, which is calculated with Levene’s test. Results from the Levene-test confirmed this assumption for the present calculations. In the 4\textsuperscript{th} semester, there was no significant difference between the students of the two study models (t-test: p=0.721). Both groups had fairly equal levels of economic content knowledge. Afterwards, students in the 6\textsuperscript{th} semester were analysed in the same way by comparing BA students and Diplom students. In

\footnotesize{\textsuperscript{10} For information on quality criteria and the factorial structure of the test items, see KUHN et al. (2014).}
the 6th semester, the means of both groups (see Tab. 2) suggested that the Diplom students might have an edge over the BA students. The results from the t-test revealed that the difference in the means of economic content knowledge was significant at a significance level of 0.09 (t-test: p=0.086). Overall, the results from the t-test were diverse. While students in the middle of their studies showed no significant difference in their economic knowledge, students at the end of their BA studies scored significantly lower than Diplom students in the same semester at least at a significance level of 0.1%.

To test the hypothesis further, a multiple linear regression analysis was conducted with the sum score of the economic content knowledge as a dependent variable. The following independent variables were included in the model: study model (BA vs. Diplom), degree course (b&e vs. business education), number of semesters, intelligence (assessed via the analogy score), school leaving grade, previous vocational training, gender, and mother tongue. The results showed that the analysed variables could explain about 25% (corr. 23%) of the variance in economic content knowledge. Furthermore, four of the independent variables were found to contribute significantly to this explained variance (see Tab. 3).

<table>
<thead>
<tr>
<th>WBT (included)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>R = 0.499</td>
<td>R² = 0.249</td>
<td>Corr R² = 0.226</td>
</tr>
<tr>
<td>Coefficient</td>
<td>B</td>
<td>Significance</td>
</tr>
<tr>
<td>(absolute term)</td>
<td>4.756</td>
<td>0.000</td>
</tr>
<tr>
<td>Study model (bachelor)</td>
<td>-0.184</td>
<td>0.449</td>
</tr>
<tr>
<td>Degree course (business education)</td>
<td>0.623</td>
<td>0.028</td>
</tr>
<tr>
<td>Semester</td>
<td>0.518</td>
<td>0.000</td>
</tr>
<tr>
<td>Gender (female)</td>
<td>-0.979</td>
<td>0.000</td>
</tr>
<tr>
<td>Analogy score</td>
<td>0.209</td>
<td>0.000</td>
</tr>
<tr>
<td>School leaving grade</td>
<td>-0.201</td>
<td>0.316</td>
</tr>
<tr>
<td>Mother tongue (German)</td>
<td>0.760</td>
<td>0.450</td>
</tr>
<tr>
<td>Vocational training</td>
<td>0.183</td>
<td>0.499</td>
</tr>
</tbody>
</table>

Tab. 3: Regression on the economic content knowledge (4th and 6th semesters)

Very important factors included verbal intelligence as well as the number of semesters. The more students had advanced in their studies, the higher they scored on the knowledge test (for longitudinal analyses, see HAPP, in prep). In this analysis, students from both the 4th and 6th semesters were taken into account, and accordingly, they had attended a different number of classes. Hence, this finding was generally in line with expectations. Further influence factors included the degree course, where business education students had an edge over b&e students, and the subjects’ gender, where male students scored higher.
The factors of previous vocational training, school leaving grade, and mother tongue did not have a significant influence on the selected sample. However, their effects already have been confirmed by statistical analyses for larger sample sizes within this project (see ZLATKIN-TROITSCHANSKAIA et al., 2013). With regard to the study model, neither the BA/MA nor the Diplom study model was found to have a significant influence on students’ content knowledge. In summary, the hypothesis could not be refuted at this stage. It would be necessary to reconsider the hypothesis within the context of all the findings. Even when further structural and personal factors were controlled, the regression analyses did not show any further significant effects associated with the study model.

4 Discussion of Results and Implications for Quality Development

The Bologna process and the associated structural and organisational changes often have been judged rather negatively in the literature (e. g., BLOSSFELD et al., 2013). For example, transparency, modularisation, the examination system, and recognition of external achievements frequently have been pointed out as areas for improvement (e. g., SANDFUCHS et al., 2011). However, as explained in the introduction, there are hardly any empirically proven results regarding the effects of the Bologna reform on study performance or learning success. In this regard, the longitudinal ILLEV study provided initial empirical findings showing an altogether diverse situation. The observed means of the students from the 6th semester showed moderate differences between both study models, with Diplom students having achieved higher test scores.

These findings do not correspond with theory, educational policy or long-term plans for the degree courses. The findings gain additional importance when they are considered alongside the available international comparative data for the Test of Economic Literacy (TEL) (see SOPER, & WALSTAD, 1987), which was the basis for the WBT, its German adaptation. GILL, & GRATTON-LAVOIE (2011) used the TEL to assess the economic knowledge of high school graduates in the U.S. upon transition into college. The subjects achieved a mean of 64 %, thus reaching almost the same level as the BA students surveyed in the present study who were at the end of their studies in their 6th semester. In comparison to the results by BUTTERS, & ASARTA (2011), this effect becomes even more striking. In their study, the high school graduates achieved a mean of 79.5 %. This finding emphasises even more the central importance of specific measures for the development of quality and for the assessment of performance in HE teaching. Particular care is advised when interpreting results and drawing potential implications for policy and practice in HE. For example, it is not yet clear whether it will be possible to replicate these results in the future or what the results may be in other faculties and HE institutions.

Nevertheless, the overall findings in this study show that quality development and performance assessments are of key importance in HE. In this context, the fact that students with a different number of semesters completed showed differences in their test scores highlights the central importance of systematic formative or pro-
cess-oriented diagnostics during university studies (for the advantages of formative assessment, see BLACK & WILLIAM, 1998). Diagnostic assessments should be introduced, if possible, at the beginning of university studies and also should take into account students’ previous knowledge. Ideally, they also should be considered in HE didactics and when teaching and learning opportunities are designed for a potentially high internal differentiation.

In times of the Bologna reform, it is paramount that the assessment of output criteria, such as content knowledge, be complemented by an assessment of outcome-oriented indicators, such as graduates’ transition into the labour market. Outcome-oriented indicators also should be generally considered in external or work-related validation of output-based results. Furthermore, international comparative findings suggest that more attention should be paid to already established systems of quality development and quality assurance. In the U.K., for example, performance management systems were introduced in HE already at the beginning of the 1990s (BROADBENT, 2007). Performance management encompasses “concepts which aim to enhance the use, and to improve the performance and competitiveness, of organisations [HE institutions] by applying novel management systems” (cf. KRAUSE, 2005). In this sense, performance management is considered a subarea of quality management (FORRESTER, 2011). In connection with performance management, the term performance refers to the degree of achieving an aim or of accomplishing a potential achievement within an organisation [HE institution] (KRAUSE, 2005). In contrast to quality management, the term performance management is closely associated with output-related and outcome-related criteria, as well as with transparency and accountability (FORRESTER, 2011).

In Germany, fewer than 10 % of universities and universities of applied sciences have introduced concepts of performance management, such as the balanced scorecard (German Rectors’ Conference 2010). However, performance management can be beneficial when applied in the HE sector, where aspects of economic viability are as important as non-financial output, such as study achievement. Performance management can support HE institutions in achieving as efficiently and effectively as possible intended aims of quality development and quality assurance, as well as in presenting their achievements externally and making them comparable, which reinforces their accountability. Measures of performance management can be applied at various levels of HE, including the macro level of HE as a whole, the individual HE institution level, the department level, the degree course level, the class level, and the micro level of the individual, such as a lecturer, a student, and so on. International experiences illustrate that both monetary and non-monetary performance criteria are highly relevant in performance management at all levels of HE. Overall, the internationally established models for quality assurance in HE provide numerous interesting suggestions on how to optimise quality development and performance assessment in HE in Germany. In addition, they present empirical research in HE with huge theoretical and methodological challenges, which have yet to be resolved.

---

11 Accreditation practice in Germany and the large amount of criticism with regard to this approach cannot be discussed in greater depth in this paper.
5 References


Authors

Dipl.-Hdl. Roland HAPP ǁ Johannes Gutenberg-Universität Mainz ǁ Jakob Welder-Weg 9, D-55099 Mainz
www.wipaed.uni-mainz.de/ls/975.php
roland.happ@uni-mainz.de

Prof. Olga ZLATKIN-TROITSCHANSKAIA ǁ Johannes Gutenberg-Universität Mainz ǁ Jakob Welder-Weg 9, D-55099 Mainz
www.wipaed.uni-mainz.de/ls/97.php
lsTroitschanskaia@uni-mainz.de