Using Vignette Testing to Measure Student Science Teachers’ Professional Competencies

Dorothee Brovelli1,*, Katrin Bölsterli1, Markus Rehm2, Markus Wilhelm1

1Science Education, University of Teacher Education Lucerne, Switzerland
2Department of Chemistry & Didactics, University of Education Heidelberg, Germany
*Corresponding author: dorothee.brovelli@phlu.ch

Received May 12, 2014; Revised July 10, 2014; Accepted July 14, 2014

Abstract Instruments to measure teachers’ professional competencies with good context validity have been discussed before. We propose a measuring instrument based on vignette testing that analyzes professional competencies for science education using extended teaching contexts. It assesses pedagogical content knowledge (PCK) and compares it to pedagogical knowledge (PK). The test consists of text vignettes that describe authentically complex teaching situations in biology, chemistry and physics and then asks open-ended questions. It is analyzed by means of content analysis, thereby integrating quantitative and qualitative methods. The evaluation of the instrument is indicative of a valid measurement instrument and has shown promising potential for assessing competencies. We used the vignette test to compare the competencies of students in different training programs and with different educational backgrounds with the objective of improving science teacher education. The testing format developed in this study proved to be able to assess teachers’ professional competencies validly in their situational context of teaching. Results show a significant increase in competencies among science teacher students in the course of their training. The results of the control group (science students not pursuing a career in teaching) are remarkably poorer. Moreover, differences were identified when comparing the vignette based competency profiles of science teacher students and teacher students of other subjects, the latter scoring significantly lower in PCK. In addition, we compared competencies of students from different teacher training programs and found differences for students trained in interdisciplinary approaches (combining the disciplines biology, chemistry and physics) and students in separate-discipline approaches.

Keywords: vignette test, professional competencies, science education, teacher training


1. Introduction

Ever since the assessment of teachers’ professional competencies became a focus of research, adequate methods to capture such competencies in a valid way have been discussed. Taking existing vignette tests into account, we describe the further development of this approach. It aims at providing an instrument for testing competencies for science teaching validly and in their situational context. Moreover it compares the competencies of students in different training programs and with different educational backgrounds with the objective of improving science teacher education.

There is a growing interest in assessing teachers’ professional competencies as a crucial determinant for effective teaching. Underlying theoretical models are predominantly inspired by Shulman [7], distinguishing pedagogical knowledge (PK), pedagogical content knowledge (PCK) and content knowledge (CK). Apart from these knowledge domains, Baumert, Kunter, Blum, Brunner, Voss, Jordan et al. [1] also consider self-regulation (SR), beliefs and motivational orientations in their study on teacher competence for mathematics instruction.

While traditional multiple-choice tests to assess teachers’ knowledge ensure “stable psychometric properties, easy test administration, and low scoring costs” [4], their de-contextualized format might not be able to reflect the complexity of teaching. An alternative are text-based or video vignettes, i.e. contextualized descriptions of classroom instruction. By measuring teachers’ or students’ ability to analyse and interpret the depicted teaching situation, their own teaching competence can be elicited and hence assessed [4,6]. To preserve the teaching context's authentic complexity, either video vignettes or comprehensive written descriptions that require different aspects of teachers’ competence are most appropriate. Video vignettes provide a maximum of situativity and authenticity and are therefore well suited to assess pedagogical knowledge [10]. However, text vignettes have the advantage of permitting summaries of longer passages of a lesson. That way, situations can be described whose interpretation requires pedagogical content knowledge like for instance a
teaching method that does not fit the content taught or a practical activity that is not well linked to scientific ideas (see [8] for the assessment of PCK in physics teaching).

Even though vignettes with a close-ended answer format can easily be analysed statistically, this format has a number of shortcomings: Stecher, Le, Ryan, Robyn und Lookwood [9] observe a tendency towards socially desirable responding. Besides, providing answer alternatives might give undesired hints on possible answers. Jeffries and Maeder [3] point out that vignettes with open-ended answer formats are appropriate to provide evidence of problem solving and critical thinking. To analyse the texts produced by the participants of the study, content analysis according to Mayring [5] is particularly suited since it represents an empirical, methodologically controlled approach and integrates quantitative and qualitative methods.

Vignette testing therefore seems a promising approach to investigating a number of questions about the development of science teachers’ competence, including the effects of different accesses into science teaching. In particular, teacher candidates can be trained in interdisciplinary or separate-discipline (biology, chemistry or physics) approaches. Moreover, many science teachers are not trained in science at the university level, but in other school subjects, and some are scientists with hardly any training in teaching.

Hence, the present study sets out to develop an instrument to assess science student teachers’ professional competencies with good context validity and to use this vignette test on assessing the teaching competence of students with different educational backgrounds.

2. Materials and Methods

A written vignette test was developed that uses authentically complex teaching situations along with open-ended questions. It mainly assesses pedagogical content knowledge (PCK) and compares it to pedagogical knowledge (PK), including a limited number of items covering content knowledge (CK) and self-regulation (SR). The proposed instrument describes eight extended teaching situations in biology, chemistry, physics and integrated science in text-format. They contain 44 problems, i.e. items that cover problematic teaching practices. Each vignette includes part of a lesson in science education. The teaching contents are chosen for authenticity and coverage of professional competencies that are generally agreed on to be necessary for successful science teaching in literature [1,8]. Examples of the teaching situations have been published elsewhere [2]. For each vignette, students had a given amount of time (5 to 10 min maximum) to write down their answers. It took approximately 70 min to complete the test, which included vignettes and demographics.

The student teachers were asked to write down a feedback on the teaching situation. Using a coding manual, depth and breadth of the students’ answers were rated by means of content analysis [5]. Four raters were trained based on the coding manual with typical answer examples. Borderline cases were discussed, and inter-rater correlation was satisfactory (Cohen’s kappa = .75, n = 152). Ratings per item were coded “0” if a participant stated no appropriate answer, “1” if a participant gave the appropriate answer, and “2” if a participant suggested a correct improvement.

Various attempts were made to ensure a high internal and external validity. Quality criteria required for mixed-methods approaches [5] were met, including communicative validation, construct validity and context validity. Moreover, several psychometric quality checks were carried out like multidimensional scaling and a subsequent confirmatory factor analysis using mean ranks scales.

Nonparametric tests were employed on the resulting ordinaly scaled data. For comparison of two groups, Mann-Whitney-U-tests were employed and for comparison of more than two groups Kruskal-Wallis-tests were used. Post-hoc analyses were Bonferroni-corrected.

386 students from three universities in Germany and Switzerland participated in the study. 67% of the participants were female. Participants were 17 to 46 years old (mean = 24.64, SD ± 4.02). It is assumed that there are no culture-specific differences between Swiss and German science education students relevant to our study. The sample consisted of students from 1) a teaching program based on a combined form of science teaching (integrating the three disciplines), with subsamples of students at different stages in the course of their training,

2) a teaching program teaching the sciences (biology, chemistry or physics) in separate disciplines, with subsamples of students at the end of their university training and after a year of practical training called “Referendariat”.

3) a university science program not aimed at pursuing a career in teaching and

4) teaching programs for subjects other than science.

Note that the students in the separate science program study in two phases: At first they attend courses at the university followed by a second phase of practical training called “Referendariat”. During that phase they are allocated to schools in order to observe other teachers’ instruction and gradually start teaching their lessons independently and also attend some courses. This does not exist in the combined science program where practical training is part of the university program taking place during the time there.

3. Results

Because the subgroup of students in the combined-science training program is rather homogenous, sharing all the rest of their courses and taking their classes in a given order, this group was selected for an investigation of the development of competencies over the course of their training. Figure 1 shows the mean ranks for the PCK- and PK-scales for these students at the beginning of their training and after 4, 6 and 8 semesters. The CK- and SR-scales are represented by broken lines because of their limited relevance due to the small number of corresponding competencies covered by the test and are not considered further. When taking all competencies into account, the Mann-Whitney-U-test revealed a significant increase with small effect size (Z=−1.89, p=.03, r=.21). A Kruskal-Wallis-test showed significant differences in PCK scales.
American Journal of Educational Research

(χ²(3)=21.31, p<0.001) and PK (χ²(3)=6.43, p<0.046) in the course of training. Bonferroni-corrected post-hoc analysis revealed an increase in PCK from start to after 8 semesters (Z=−4.996, p<0.001, r=0.562) with large effect size, and an increase in PK from start to after 4 semesters (Z=−2.394, p<0.017, r=0.323) with a subsequent decrease from after 4 to after 6 semesters (Z=−2.355, p=0.019, r=0.279).

Figure 1. Mean ranks for combined-science teacher students in the course of training for the SR, PK, PCK and CK scales

The measured competencies of the separate-science students – whose training program allows for taking the classes in a random order – do not show a similar trend. However, combined-science teacher students right before graduation achieve significantly higher results in PCK (Z=−2.925, p=0.003, r=0.27) than separate-science teacher students (see Figure 2). In PK they do not score significantly different (p=0.471).

Figure 2. Mean ranks for science teacher students in combined and separate sciences programs for the PCK, PK, CK and SR scales

In Figure 3, the mean ranks for combined science teacher students are compared to those for science students of a technical university not enrolled in any teacher training programs. Both groups had taken the same amount of science courses. Significant differences become evident when taking all competencies into account (Z=−1.653, p=0.049, r=0.21), the teacher students scoring significantly higher on PK and PCK with large effect sizes (PK: Z=−3.924, p<0.001, r=0.50, PCK: −3.596, p<0.001, r=0.46).

Figure 4 shows the mean ranks for teacher students of school subjects other than science and compares them to the mean ranks for a subsample of the science teacher students (chosen for comparable overall university training time). While not scoring significantly different when taking all competencies into account (Z=−1.478, p=0.139, r=0.13), the science teacher students achieve significantly higher results in PCK (Z=−2.301, p<0.01, r=0.21) and lower results in PK (Z=−2.843, p=0.004, r=0.26) when compared to teacher students of other subjects.

4. Discussion and Conclusions

The testing format developed in this study proved to be able to assess teachers’ professional competencies validly in their situational context of teaching. Results show a significant increase in competencies among science teacher students in the course of their training and allow for the identification of shortcomings in PCK and PK by analysing the subcategories mainly causing the differences [2]. The data permits a number of conclusions when comparing different groups of students. The results of the group of science students not pursuing a career in teaching are remarkably poorer identifying specific needs of career switchers into science teaching. Moreover, differences were identified in the competency profiles of science teacher students and teacher students of other subjects. The latter compensate their shortcomings in PCK and CK by focusing on PK with implications for the teaching style of teachers coming from other subjects. Differences for students from different teacher training programs (separate vs. combined disciplines) show potential for improvement of the programs, but also confirm that the newly developed combined science teaching program is successful in developing students’ PCK, resulting in equal
or higher competencies than for the separate science program.

References


