2.5 Development of Prospective Physics Teachers’ Professional Knowledge and Skills during a One-Semester School Internship

Vogelsang, C., Borowski, A., Kugelmeyer, C., Riese, J., Buschhüter, D., Enkrott, P., Kempin, M., Reinhold, P., Schecker, H., and Schröder, J.

Abstract

In academic teacher education programs prospective physics teachers are supposed to acquire professional knowledge and skills that enable them to carry out effective instruction. However, it is unclear which knowledge has an impact on teaching quality and how knowledge and teaching skills develop throughout studies. In particular, this is the case for practical parts of teacher education programs like school internships or other forms of field experiences. Therefore, we examine the development of pre-service physics teachers‘professional knowledge over a one-semester internship in schools. Furthermore, we analyze the development of their skills in (a) planning physics lessons, (b) reflecting physics lessons, and (c) explaining physics over that internship using an innovative approach of standardized performance assessments. So far, our longitudinal analyses of a cohort of prospective physics teachers from four German universities hardly show any development of their professional knowledge and skills during the internship. Further analyses are needed to gain more insight into this low efficacy of the internship.

O. Zlatkin-Troitschanskaia et al. (Hrsg.), *Student Learning in German Higher Education* https://doi.org/10.1007/978-3-658-27886-1_6

Keywords

Science education, pre-service teacher education, professional knowledge, teachers' professional development, performance assessment, pre-post-measurement, field experience

1 Introduction

The primary goal of teacher education is to support future teachers in developing professional knowledge and skills to meet the challenges of their profession. Common models treat professional knowledge as a key component of professional competence, whereas competence can be described as 'the latent cognitive and affective-motivational underpinning of domain-specific performance in varying situations' (Blömeke et al. 2015, p. 3). Therefore, a great amount of teacher education programs focus on the development of such knowledge, especially in the German teacher education system. In Germany, teacher education consists of three consecutive phases necessary to become a teacher (Cortina and Thames 2013). In the first phase, future teachers enrol in a teacher education study program at a university, including a three-year bachelor's degree followed by a two-year master's degree. The second phase consists of a 12 to 24 months in-school induction program. While the first phase focuses on the acquisition and development of theoretical knowledge, the second phase emphasizes practical teacher training. The third phase aims at the further professional development of in-service teachers. The underlying model of these phases can be described as a functional chain (Diez 2010). During their university studies, future teachers are meant to acquire knowledge, which they have to apply in their teacher training course afterwards. In this perspective, it is assumed that teachers use their professional knowledge as a resource to perform their daily tasks. This model poses some challenges for teacher educators creating teacher education programs. They have to ensure that prospective teachers acquire knowledge, which is relevant for teaching. It has to be part of the knowledge base for teaching (van Driel et al. 2001), which allows teachers to develop skills to carry out high-quality instruction. To meet this challenge, most federal states in Germany implemented a one-semester internship at a school as part of their master's degree programs for teachers (practical semester). It is meant to enable the use of theoretical knowledge and to gather first teaching experiences already before the second phase. During their one-semester internship, in addition to teacher training by expert teachers at school, all participating students also attend supporting courses at the university (usually one day of the week). However, it remains an open question to which extent academic teacher education programs contribute to the development of professional knowledge and skills and which types of knowledge have an impact on the quality of performance in teaching situations. This is especially the case for field experiences in teacher education like long-term internships at schools.

In our research project Profile-P+ (*Professional Knowledge in Academic Physics Teacher Education*), these questions are addressed. We focus on prospective physics teachers for secondary schools in Germany. We examine the development of their professional knowledge by longitudinal studies over the first two years of their bachelor's degree program and evaluate their development of professional knowledge and skills to cope with three typical requirements for physics teachers (planning physics lessons, reflecting on physics teaching, explaining physics) in a longitudinal section during a practical semester. Moreover, we analyze the relationship between professional knowledge and skills. In this chapter, we present preliminary results regarding professional development during the one-semester internship.

2 Theoretical Background

2.1 Physics Teachers' Professional Knowledge

Blömeke, Gustafson and Shavelson (2015) describe competence as a continuum that regards cognitive and affective-motivational dispositions as the basis for situation-specific skills, which in turn enable performance in complex real-life situations. Professional knowledge is seen as one key-part of these dispositions of the professional competence for teaching. Following recent work based on the influential considerations of Shulman (1986), we focus on three dimensions of professional knowledge: content knowledge, pedagogical content knowledge and general pedagogical knowledge. Content knowledge (CK), also known as subject matter knowledge, comprises knowledge of the contents and methods of the subject taught as well as epistemological aspects. For the domain of physics, there are different models of content knowledge (Woitkowski and Borowski 2017), which usually differentiate the "depth" of knowledge according to curricular levels (school knowledge, university knowledge). We extend this structure by a specific form of CK assumed to be a specific basis for teaching: deeper school knowledge. In our approach, deeper school knowledge refers to a meta-perspective on school knowledge, for instance, identifying suitable problem solving strategies for specific physics problems or the discussing structural relationships between physics concepts. Pedagogical content knowledge (PCK) refers to knowledge that teachers need to prepare and structure content in a way that is appropriate for their students. For the domain of science (especially physics), a large number of models have been proposed, which have recently been brought together in the Refined Consensus model of PCK (Carlson et al. 2019). For example, PCK for physics teaching contains knowledge about students' alternative ideas of physics concepts, about instructional strategies and about implementing experiments in physics instruction. Finally, pedagogical knowledge (PK) contains knowledge about learning principles and pedagogical concepts like knowledge about classroom management or motivation (e.g. König et al. 2011). Surely, CK, PCK, and PK do not cover the full complexity of teachers' professional knowledge; however, these three dimensions are mirrored in the typical structure of academic physics teachers education in Germany. During their studies, student physics teachers take courses focusing on content knowledge (CK), general pedagogy (PK) and also on concepts of physics education (PCK).

2.2 The Influence of Physics Teachers' Professional Knowledge on Their Teaching Skills

The professional skills of physics teachers have been assessed by the quality of their instruction (Diez 2010). However, little correlation has been observed between physics-related professional knowledge and the quality of physics teaching or student learning in physics lessons. Cauet et al. (2015) could not find any correlation between physics teachers' professional knowledge (CK and PCK measured by written tests containing multiple choice and open questions), students learning (measured by written tests containing multiple choice and open questions) and the level of cognitive activation of the observed physics lessons (measured by video-analysis) in German secondary schools. Similar results are reported by Ohle, Boone & Fischer (2014). They analyzed the impact of K4-elementary teachers' CK of a specific physics topic on their students' achievement (both measured by multiple choice tests). „Results showed that neither teachers' interest nor content knowledge impacted students' outcomes directly." (Ohle et al. 2014, p. 14).

These results are independent of how quality of teaching was measured, for instance, by high inference ratings of instruction (e.g. Korneck et al. 2017) or by analyzing the content structure of observed lessons (Liepertz and Borowski 2018). In the sense of the Refined Consensus Model, hardly any connections between enacted PCK and personal or collective PCK could be observed (Carlson et al. 2019). This might indicate a lack of relevance of the assessed knowledge for teaching.

However, in all of these studies, the assessment instruments used show good curricular validity often based on expert ratings of teacher educators or experienced physics teachers.

Another reason for the missing link between teaching quality and students' achievement might result from the complexity of real teaching situations. Teaching in real classrooms is affected by many heterogeneous context factors that might suppress existing relations between teachers‘ knowledge and their actual teaching performance (Kulgemeyer and Riese 2018). To obtain a clearer picture, the quality of teaching should, therefore, be researched in a more standardized way. This can be achieved according to concepts from medicine education (Miller 1990). So-called 'Objective Structured Clinical Examinations (OSCE)' simulate typical standard situations for health professionals, often with the help of trained actors. These scenarios are designed as authentic as possible to mirror real situations occurring in everyday professional life. At the same time, they are standardised in such a way that a high degree of comparability is guaranteed. Scoring sheets can be used to achieve high test standards (Walters et al. 2005). Kulgemeyer and Riese (2018) developed such a performance assessment for explaining physics in the form of a role-playing situation. They examined the correlation between prospective physics teachers‘ professional knowledge (CK and PCK, measured by written tests containing multiple choice and open questions) and their skills in explaining physics (measured by performance assessment). The results showed significant correlations between explaining skills and professional knowledge (for details, see Kulgemeyer and Riese 2018, p. 18).

2.3 Development of Professional Knowledge and Skills during Field Experiences

School internships or field experiences are an integral component of typical teacher education programs. Internships have multiple aims, for example, to provide opportunities for prospective teachers to make first teaching experiences on their own, to acquire basic skills for teaching (like planning a lesson or reflect on one's teaching) and to cope with the gap between academic learning and the demands of professional practice in the field (Cohen et al. 2013). However, it is unclear, whether internships contribute to those aims. Most research on the effectiveness of internships or field experiences, especially in the context of German teacher education, relies on student teachers‘ self-reports like self-rated skills or self-rated instructional quality (Besa and Büdcher 2014). Self-reports are often suspected of bias, like the tendency of participants to rate themselves in a favourably way. However, studies hardly used more proximal assessments to evaluate the development of skills during internships. Holtz and Gnambs (2017) examined the change of instructional quality of student teachers throughout a 15-week school internship, measured by ratings of two observed lessons by different groups, expert teachers, students, and self-ratings. The results showed an improvement of instructional quality, but they are related to many different school subjects not specifically to physics. Volmer et. al (2019) investigated the skills of prospective elementary teachers during a one-semester internship using open written reflections on a given videotaped lesson. They report a significant increase in reflection quality between before and after the internship.

3 Research Questions

Prior studies showed few and inconclusive results regarding the development of prospective physics teachers' professional knowledge during academic teacher education and the relation between knowledge and skills. To contribute to these research gaps, we address the following overarching research questions (RQ) in the project Profile-P+:

1. How does the professional knowledge (CK, PCK, PK) of prospective physics teachers develop over a one-semester internship (pre, post)?
2. How do the skills of *planning* physics lessons, *explaining* physics and *reflecting* on physics lessons develop over a one-semester internship (pre, post)?

In our study, we collect data of additional personal characteristics adjusted for mathematical skills, attitudes to explanation, learning opportunities, and demographic data.

Since there are hardly any opportunities for systematically learning more theoretical knowledge over a one-semester internship, we expect little to no increase in professional knowledge (RQ1). However, we expect substantial increases in all three analyzed skills (RQ2) as the internship should provide many opportunities for lesson planning and reflecting and explaining physics in classroom instruction in an authentic school setting.

4 Design and Sample

We used a longitudinal approach to follow cohorts of prospective physics teachers in master's degree programs at four german universities in three different federal states.

Although the internships vary in some details between the four participating universities, the overall structure is similar. During the internships – lasting about five months – the prospective physics teachers receive guidance in their teaching from experienced teacher mentors at a school four days a week. This part of the internship is supposed to provide practical insights into German teachers' daily routines and to enable first self-sufficient teaching experiences. The mentors are supposed to provide feedback on the instruction carried out by the teacher students and to help them reflect on their teaching. One day a week, the prospective physics teachers visit university courses (one for each of the two teaching subjects, one for general pedagogics). Part of these courses is a reflection on experiences made at the internship schools taking based on theoretical concepts of PCK and PK. Also, in these courses prospective physics teachers are supported in planning their lessons. This structure (in-school-training combined with university courses) is typical for one-semester internships in teacher education programs of the most federal states in Germany.

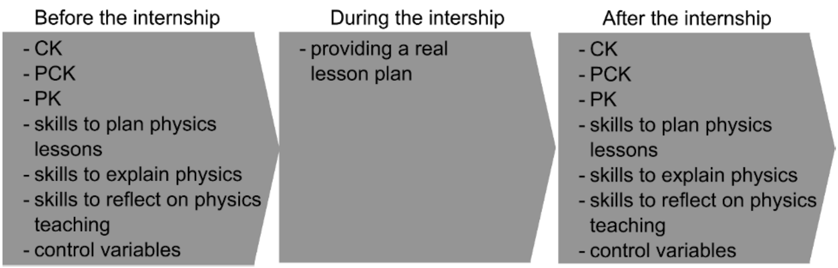

Figure 1 Study design: longitudinal section during the practical semester in the master's program

For our study, the professional knowledge (CK, PCK, PK), as well as the skills to plan physics lessons, to explain physics, and to reflect on physics teaching, were assessed before and after completing the one-semester internship (Figure 1). The extensive data collection required five test dates before and after the internship each, which were realised within the accompanying university courses taken by the students. Tests were scheduled on different days for each measurement point.

The total test time per person and measurement point was 330 minutes. The complete sample consists of $N = 80$ prospective physics teachers. On average, they were in the ninth semester of their study programmes (semester: $M = 9.03$; $SD = 3.03$) during the internship and 25 years old (age: $M = 25.41$; $SD = 5.09$). 61 % of the students are male, 39 % female.

The data of all participating universities has been pooled for analyses. Cohort effects have been checked. Due to the high testing time and the facts that the participation was voluntary not all teacher students took part in all the assessments. Therefore, we used different subsamples for analyses depending on the construct we focused on. To control for possible effects of decreasing test motivation we used a short test-motivation scale after every performance assessments (for details, see Vogelsang et al. 2019). Dependent t-tests were carried out to analyse changes of professional knowledge and performance during the internship (pre, post).

5 Instruments

For data collection, we used two widely tested and comprehensively validated instruments from our previous project Profile-P (written test for PCK, performance assessment for the skills of explaining physics) (Riese et al. 2015). The instrument for the assessment of CK was adopted from Profile-P and further developed with greater emphasis on deeper school knowledge. The performance assessments for the skills of planning physics lessons and reflecting on physics lessons have been newly developed for this study. All the instruments refer to the physical content of mechanics to establish the necessary content comparability.

5.1 Content Knowledge

CK is seen as a major component of teachers' professional knowledge. In line with previous research we assume, that physics teacher need specific forms of CK, which is of special importance to plan high-quality physics lesson or to reflect on their teaching properly. Following already developed models (Kirschner 2013), we distinguish between three dimensions of CK: school-, university- and deeper school knowledge (SK, UK, DSK). School knowledge is described by the official school curricula and university knowledge can be operationalized by the university curriculum. Based on the approach of Riese et. al. (2015), we defined deeper school knowledge as (1) identifying relations between physics concepts, (2) handling model limitations, and (3) identifying suitable problem-solving approaches.

As described by Riese and Reinhold (2012), this knowledge is assumed to be of special importance for teachers.

Based on this model, we developed a written test consisting of 48 single-choice-items focusing on different aspects of mechanics (velocity and acceleration, Newtons Laws, Conservation of Energy). The total test time is 50 minutes. Curricular validity was ensured by analyses of typical school and university textbooks, by analyses of school curricula and analyses of the physics curricula of the participating universities. Stimulated recall interviews with N = 8 physics teacher students after taking the test indicate, that the items are perceived as appropriate tasks in academic teacher education. For construct validity, we investigated whether the structure of our model corresponds to the structure of the empirical data. We compared Rasch models with different dimensions based on data of N = 861 physics teacher students in a bachelor or master program at twelve German universities combined with students studying the research-oriented physics bachelor program. The results show that a three-dimensional model is to be preferred over two- or one-dimensional models. For all the sub-scales, satisfactory reliability was found (EAP-reliability ranging from 0.76 to 0.84; for details, see Vogelsang et al. 2019, p. 484).

5.2 Pedagogical Content Knowledge

For the assessment of PCK, another key component of physics teachers' professional knowledge, we used a written instrument already developed in our previous project Profile-P (Riese et al. 2015). The underlying comprehensive model of physics teachers' PCK was developed using different conceptualizations of PCK in science subjects (e.g. Lee & Luft 2008; Magnusson, Krajcik et al. 1999) and also considering curricula analyses. The test instrument focuses on four aspects of PCK: experiments, instructional strategies, students' misconceptions and how to deal with them and physics education concepts like conceptual change. It includes open situational judgment items as well as complex multiple-choice items (multiple select, 43 items). The total test time is 65 minutes. To ensure validity, several steps of validation were taken in our previous project Profile-P (Riese et al. 2015). We investigated content validity by analyzing curricula and expert ratings from educators at four universities. Also, a think-aloud study with $N = 15$ prospective physics teachers was carried out to check, if items could be solved using only CK rather then PCK (Gramzow et al. 2013). For construct validity, one-dimensional and four-dimensional Rasch models were compared, indicating a better matching of the data by a four-dimensional model (for details, see Kulgemeyer and Riese

2018). For all the sub-scales, satisfactory moderate reliability was found in different studies (e.g. Kulgemeyer and Riese 2018; Riese et al. 2015). For the rather small sample of this study we used manifest scores and found a rather low but sufficient reliability for the total score (*Cronbach's* $\alpha = 0.66$).

5.3 Pedagogical Knowledge

For the assessment of pedagogical knowledge, we used an adapted short-version (Riese and Reinhold 2012) of a written instrument from Seifert and Schaper (2012). The full version was used in several studies in Germany (Mertens and Gräsel 2018), also the short-version (Riese and Reinhold 2012). The short-version addresses two aspects of PK: general instructional strategies and classroom management. The total test time is 15 minutes. Since we focus on the subject-specific components of professional knowledge in Profile-P+, no further validations were carried out by ourselves. We rely on the results from previous research where sufficient arguments for construct validity and content validity regarding teacher education in Germany has been reported (for details, see Seifert and Schaper 2012).

5.4 Skills of Planning Physics Lessons

The process of teachers'lesson planning can be described as a recursive process. To plan a lesson, teachers have to analyze preconditions, plan certain classroom actions and reflect on their planning decisions (Shavelson and Stern 1981). Experienced teachers mostly do not elaborate on their lesson plans, as they have scripts and routines to fall back on (Stender 2014). However, beginning or student teachers need to develop such scripts. Therefore, they need to plan actual lessons. In doing so, it is assumed, that student teachers heavily rely on their – more theoretical – professional knowledge. To analyze this assumption, we developed a performance assessment for prospective physics teachers'skills to plan physics lessons. We followed Miller's (1990) approach and developed an instrument to assess planning skills in a standardized performance situation, in which student teachers need to plan a whole lesson instead of reproducing knowledge about lesson planning. The paper-pencil instrument puts students in a situation where they have to plan a lesson about Newton's third law. Therefore, a short description of the class and their learning prerequisites is provided and specific learning objectives are set. The lesson plan has to be documented on a prestructured planning paper, which suggests some mandatory parts of physics lessons to allow higher comparability.

To evaluate the quality of lesson plans, we developed a theoretical model, which contains different aspects of physics lesson planning (e.g. implementation of experiments, exercises, contexts, learning objectives and preconditions). The model was developed by using subject-specific literature about lesson planning and it was combined with an inductive approach, resulting in a codebook with currently $N = 59$ coding items. By using the codebook, we have so far coded $N = 141$ out of 160 individual lesson plans, resulting in 66 sets of pre and post data. For interrater agreement, $N = 52$ of the lesson plans were double coded. The agreement amounts to 89 %, Gwet's AC 1 is .849, which indicates good agreement. We carried out multiple steps of validation. Interviews with three teacher trainers were conducted and their judgements regarding the quality of three selected lesson plans were compared to corresponding results provided by the codebook. The results indicate an agreement among the teacher trainers about the perceived order of quality of those lesson plans as well as an agreement of their mean grades' order to the order of the planning score provided by the codebook. The standardized lesson plans were also compared to real lesson plans, created during the internship by the same students. We found an indication of similar planning behaviour comparing the assessment and the real lesson plans. Based on all 59 coding-items a sum score was built. For the sample of this study, the scale-reliability of the total planning score is sufficient (*Cronbach's* $\alpha = 0.79$).

5.5 Skills of Explaining Physics

To assess the skills to explain physics, we used a performance assessment already developed in our previous project Profile-P (Riese et al. 2015). Explaining is core part of physics instruction and therefore, a central skill of physics teachers (Geelan 2012). Explanations in the context of instruction can be described as a dialogic process, in which the teacher tries to communicate a scientific concept to one or more students. In this process, the teacher has to consider two aspects. First, the explanation has to represent the scientific concept in an adequate way, for example, its structure, highlighting the major aspects (subject-adequate). Second, the teacher has to consider the students' needs, for example, considering their supposed prior knowledge or any misconceptions (addressee-oriented). During the process, the teacher also hat to evaluate his explanation, for instance, by asking questions to the student. We developed a dialogic explaining performance assessment simulating an authentic face-to-face explaining situation. After a short preparation time using standardized materials, participants had to explain a topic of mechanics to a student. The explaining attempts were videotaped. The student has been trained to behave in a standardized way during the explaining situation (e.g. giving specific prompts as feedback). All videotaped explanations were analyzed using the model of explaining physics by Kulgemeyer and Tomczyszyn (2015). This model distinguishes between the two aspects of explanation quality (subject-orientation, addressee-orientation) represented by 12 categories for appropriate resp. inappropriate explaining, for instance, explaining physics concepts in everyday language. Based on the codings of the videotaped explanations, an explaining performance Index (PI) was build. For this assessment, also several steps of validation have been conducted (Kulgemeyer and Riese 2018). The PI predicted expert decisions on for the better explaining quality when a pair of videos was compared to a moderate to great extent. In addition, expert interviews have been carried out to ensure content validity. Interrater-reliability of two independent raters reached accordance of 91 %. In previous studies, also suitable reliability of the PI has been reported (for details, see Kulgemeyer & Riese, 2018). For the sample of this study, the scale-reliability of the PI is sufficient (*Cornbach's* $\alpha = 0.77$).

5.6 Skills of Reflecting on Physics Lessons

Reflection can be described as a spontaneous, common, real thinking process that gives coherence to an initially incoherent and unclear situation (Clarà 2015). Plöger, Scholl and Seifert (2015) developed a multi-stage model for reflection on lessons, which was adapted for reflecting physics lessons by modifying specific challenges in the fields of CK and PCK (Figure 3). The model distinguishes three dimensions of reflection. The elements of reflection includes four steps of reflection: (1) description of the framework conditions of a lesson (e.g. students' pre-knowledge, teaching goals) and the teaching situation, (2) evaluation of the described teaching situation, (3) providing alternatives for the observed behaviour and (4) drawing of consequences for (a) the following lesson, (b) the development of the reflective individual, or (c), in case of reflecting others' actions, the development of the teacher observed (Nowak et al. 2019). The reasoning dimension represents another aspect of reflection quality. Reflection quality is higher, if evaluations, alternatives and consequences are reasoned rather than given as spontaneous subjective judgements. The last dimension indicates which knowledge base (CK, PCK, PK) is subject in the point of reflection addressed. For example, reflecting on how to deal with students' misconceptions would be based on PCK.

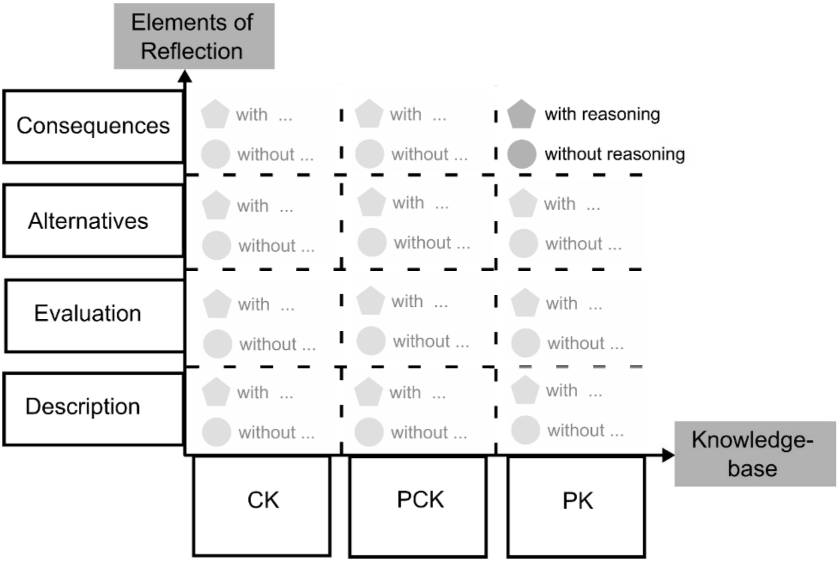

Figure 2 Model of reflections on physics lessons

Following Miller's (1990) approach, we developed a performance assessment for the skills of prospective physics teachers to reflect on physics lessons. We conduct an interactive online test simulating a situation, in which the student teachers have to reflect on the video-taped physics lesson of a fictive fellow intern. The fictive intern (Robert) asks the test person in a simulated dialogue for advice. The videotaped physics lesson focuses on Newton's third law and the conservation of momentum. It was created based on a script by persons acting as school students. The fictive lesson includes standard parts of physics instruction (e.g. an experiment) and typical problems of teaching physics (e.g. misconceptions) that provide reflection causes (RC). These RCs were assigned to the knowledge bases of the reflection model by an expert rating. During the test, the student teachers receive information about the context and they give Robert feedback on particular parts of the lessons. The verbal answers are recorded and categorised using qualitative content analysis by Mayring (2015).

The coding categories are used to evaluate the quality of reflection, for example, whether or not they give a reasoned evaluation or present alternatives for one RC. The sum of coded categories is assumed to be a measure for the skills to reflect on physics teaching. We formed a score for general reflection skills (RS). $N = 32$ complete data-sets (pre and post) have already been analyzed. Interrater agreement has been determined by triple-coding of 17 % of the performance assessments reaching an average of .882 for Gwet's AC1. That indicates a very good agreement. For the small (sub)sample of this study the scale reliability of the reflection score is rather low (*Cronbach's* α = 0.46) before resp. after (*Cronbach's* α = 0.52) the internship so far.

6 Results

Data of between 20 and 80 physics teacher students has been analyzed so far, depending on the particular construct. In this chapter, we present preliminary results.

Tables 1 and 2 describe the results regarding the development of professional knowledge differentiated into subscales for each test. CK was scaled using a three-dimensional Rasch model (for the scoring, see Section 5.1). Therefore, latent EAP-scores are reported. For PCK and PK manifest sum-scores are reported. All analyses result from dependent t-tests for particular knowledge scores.

Table 1 Development of CK (latent scores)

	N	t_1		t_2			
CK	80	M	SD	M	SD	p	d
school knowledge (SK)		0.99	0.86	1.21	0.72	0.000	0.64
university level (UK)		0.80	0.47	0.92	0.54	0.000	0.84
deeper school knowledge (DSK)		0.69	0.47	0.96	0.31	0.000	0.99

Table 2 Development of PCK and PK (sum-scores, in %)

	N	t_1		t_2			
PCK	63	M	SD	M	SD	p	d
experiments		0.50	0.20	0.55	0.23	0.106	0.21
physics education concepts		0.54	0.18	0.63	0.14	0.000	0.59
instructional strategies		0.38	0.19	0.39	0.16	0.740	0.04
students' misconceptions		0.48	0.18	0.53	0.20	0.043	0.26
PK	58	M	SD	M	SD	p	d
		0.35	0.13	0.44	0.11	0.000	0.82

Regarding CK, a significant increase in all subdimensions can be found (medium to large effect, *Cohens's d* ranging from 0.64 to 0.99, Table 1). Regarding PCK, significant increases can be seen in the sub-scales "student's misconceptions" (small effect, *Cohen's d* = 0.26, Table 2) and "physics education concepts" (medium effect, *Cohen's d* = 0.59, Table 2). Furthermore, a significant increase with a large effect occurred regarding PK (*Cohen's d* = 0.82, Table 2).

Table 3 contains the results of the performance assessments. All scores represent manifest scores (Sections 5.4 to 5.6). As Table 3 shows, only for one out of the three performance assessments a significant increase with a small effect can be observed so far (planning physics lessons, *Cohens's d* = 0.32).

Table 3 Development of skills regarding three standard teaching situations (sum-scores, in %)

		t_1		t_2			
	N	M	SD	M	SD	p	d
planning physics lessons	66	0.49	0.11	0.53	0.13	0.01	0.32
reflecting on physics teaching	37	0.11	0.03	0.12	0.04	0.404	0.16
explaining physics	20	0.30	0.15	0.33	0.12	0.420	0.24

7 Discussion

The newly developed performance assessments enable valid interpretations (Vogelsang et al. 2019) of the data reflecting three important skills of prospective physics teachers: planning and reflecting on physics lessons, and explaining physics. Regarding the one-semester internship, significant increases in all sub-dimensions of CK, in two sub-dimensions of PCK, and for PK could be observed. Due to fewer (formal) learning opportunities during the whole internship program (teacher training at schools and accompanying courses at a university) we expected little increase in more theoretical professional knowledge. However, we found medium to large effects in terms of professional knowledge. In addition, contrary to our expectations, no increases of the skills to reflect on physics teaching and to explain physics could be identified so far. That leads to the question of whether typical long-term school internships programs really contribute to acquiring professional skills, especially regarding the expected application of theoretical knowledge for high-quality performance in teaching situations (which is often expected, cf. Cohen et al. 2013). The non-significant development of skills for reflecting and explaining measured by standardized assessments may indicate a lack of effectiveness of the internship regarding central goals.

It should be noted that the results presented are preliminary and based on small sample sizes – the statistical power seems to be limited. The collected data have not been coded completely up til now and some of the coding categories of the performance assessments for planning and reflecting will be finetuned for further analyses. Although the sample is small–due to the small number of physics teacher students in Germany in general–it should be noted that we were able to carry out almost complete surveys at the participating four universities.

Other limitations lie in the low reliabilities of some sub-scales of the knowledge assessments and in the score regarding the skills to reflect on physics teaching. Besides, further improvements of the codings for the performance assessments have to be done, which might lead to an increase of the reliability and might help to identify smaller effects. For all instruments, core arguments for the construct validity will also be checked based on an extensive nomological network. Furthermore, the use of performance assessments following approaches from medical education is quite uncommon in teacher education at least in Germany (see also Kuhn et al. in this volume). One possible reason, why no effects could be observed, might lie in this uncommon testing situations for the participating students. Another reason might be a decrease in the test motivation over the internship due to the long-lasting data collection. However, such a decrease could only be identified between the measurement points of the assessment for reflection skills (based on the results of a short test-motivation scale, Vogelsang et al. 2019).

Further analyses are needed to shed more light on the effectiveness of long-term internships in teacher education. We tried to research changes in professional knowledge and skills before and after the intership using pre-post-measurements. Future studies should take a closer look at the learning processes of the student teachers *during* the internship. Although we have information regarding the content of the accompanying courses at the universities and, additionally, self-reports of the prospective teachers on their experiences with their mentors during the intership, there is a lack of information on learning processes in detail. Future studies focussing on changes in shorter time periods linked to training situations could be a useful additional approach.

Funding

Profile-P+ was funded by the KoKoHs program of the German Ministry for Education and Research (01PK15005A-D).

References

Besa, K. S., & Büdcher, M. (2014). Empirical evidence on field experiences in teacher education. In K. H. Arnold,, A. Gröschner, & T. Hascher (Eds.). (2014), *Schulpraktika in der Lehrerbildung: Theoretische Grundlagen, Konzeptionen, Prozesse und Effekte* (pp. 129–145). Waxmann.

Blömeke, S., Gustafsson, J. E. & Shavelson, R. J. (2015). Beyond dichotomies: Competence viewed as a continuum. *Zeitschrift für Psychologie*, *223*, 3–13.

Carlson, J., Daehler, K.R., Alonzo, A., Barendsen, E., Berry, A., Borowski, A., & Wilson, C. (2018). The Refined Consensus Model of pedagogical content knowledge in science education. Repositioning pedagogical content knowledge in teachers' knowledge for teaching science. In A. Hume, R. Cooper & A. Borowski (Eds.), *Repositioning pedagogical content knowledge in teachers' knowledge for teaching science* (pp. 77–92). Singapore: Springer.

Cauet, E., Liepertz, S., Kirschner, S., Borowski, A., & Fischer, H. E. (2015). Does it matter what we measure? Domain-specific professional knowledge of physics teachers. *Revue Suisse des sciences de l'education*, *37(3)*, 462–479.

Clarà, M. (2015). What Is Reflection? Looking for Clarity in an Ambiguous Notion. *Journal of Teacher Education*, *66(3)*, 261–271.

Cohen, E., Hoz, R., & Kaplan, H. (2013). The practicum in preservice teacher education: a review of empirical studies. *Teacher Education*, *24*, 345–380.

Cortina, K. S., & Thames, M. H. (2013). Teacher education in Germany. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss, & M. Neubrand (Eds.), *Cognitive activation in the mathematics classroom and professional competence of teachers* (pp. 49–62). Boston: Springer.

Diez, M. E. (2010). It Is Complicated: Unpacking the Flow of Teacher Education's Impact on Student Learning. *Journal of Teacher Education*, *61(5)*, 441–450.

Geelan, D. (2012). Teacher explanations. In B. Fraser, K. Tobin, & C. McRobbie (Eds.), *Second international handbook of science education* (pp. 987–999). Dordrecht: Springer.

Gramzow, Y., Riese, J., & Reinhold, P. (2013). Prospective physics teachers' pedagogical content knowledge – Validating a test instrument by using a think aloud study. *European Science Education Research Association 2013 Conference*, Nicosia, Cyprus.

Harden, R., Stevenson, M., & Wilson, W. G. W. (1975). Assessment of clinical competence using objective structured examination. *British Medical Journal*, *1*, 447 – 451.

Holtz, P., & Gnambs, T. (2017). The improvement of student teachers' instructional quality during a 15-week field experience: a latent multimethod change analysis. *Higher Education*, *74(4)*, 669–685.

Kirschner, S. (2013). *Modellierung und Analyse des Professionswissens von Physiklehrkräften [Modeling and Analyzing Professional Knowledge of Pre-Service Physics Teachers]*. Berlin: Logos.

König, J., Blömeke, S., Paine, L., Schmidt, W. H., & Hsieh, F. J. (2011). General pedagogical knowledge of future middle school teachers: On the complex ecology of teacher education in the United States, Germany, and Taiwan. *Journal of Teacher education*, *62(2)*, 188–201.

Korneck, F., Krüger, M. & Szogs, M. (2017). Professionswissen, Lehrerüberzeugungen und Unterrichtsqualität angehender Physiklehrkräfte unterschiedlicher Schulformen [Professional knowledge, teacher beliefs and instructional quality of prospective physics teachers of different school types]. In H. Fischler & E. Sumfleth (Eds.), *Professionelle Kompetenz von Lehrkräften der Chemie und Physik [Professional Competence of chemistry and physics teachers]* (pp. 113–133). Berlin: Logos.

Kulgemeyer, C., & Riese, J. (2018). From professional knowledge to professional performance: The impact of CK and PCK on teaching quality in explaining situations. *Journal of Research in Science Teaching*, *55(10)*, 1393–1418.

Kulgemeyer, C., & Tomczyszyn, E. (2015). Physik erklären – Messung der Erklärensfähigkeit angehender Physiklehrkräfte in einer simulierten Unterrichtssituation [Explaining physics – Measuring teacher trainees ' explaining skills using a simulated teaching setting]. *Zeitschrift für Didaktik der Naturwissenschaften*, *21(1)*, 111–126.

Lee, E., & Luft, J. A. (2008). Experienced secondary science teachers ' representation of pedagogical content knowledge. *International Journal of Science Education*, *30(10)*, 1343–1363.

Liepertz, S., & Borowski, A. (2018). Testing the Consensus Model: relationships among physics teachers' professional knowledge, interconnectedness of content structure and student achievement. *International Journal of Science Education*. doi: 10.1080/09500693.2018.1478165

Magnusson, S., Krajcik, J., & Borko, H. (1999). Nature, sources, and development of pedagogical content knowledge for science teaching. In J. Gess-Newsome & N. Lederman (Eds.), *Examining pedagogical content knowledge* (pp. 95 – 132). Dordrecht: Kluwer Academic Publishers.

Mayring, P. (2000). Qualitative Content Analysis. *Forum: Qualitative Social Research*, *1(2)*, 1–10.

Mertens, S., & Gräsel, C. (2018). Entwicklungsbereiche bildungswissenschaftlicher Kompetenzen von Lehramtsstudierenden im Praxissemester [The development of educational competences during long-term internships in teacher education]. *Zeitschrift für Erziehungswissenschaft*, *21(6)*, 1109–1133.

Miller, G.E. (1990). The Assessment of Clinical Skills/Competence/&Performance. *Academic Medicine*, *64(9)*, 63–67.

Nowak, A., Kempin, M., Kulgemeyer, C., & Borowski, A. (2019). Reflexion von Physikunterricht [Reflection of Physics Teaching]. In C. Maurer (Eds.), *Naturwissenschaftliche Bildung als Grundlage für berufliche und gesellschaftliche Teilhabe [Scientific Education as a basis for vocational and social participation]* (pp. 838–841). University of Regensburg.

Ohle, A., Boone, W. J., & Fischer, H. E. (2014). Investigating the impact of Teachers' Physics CK on Students outcomes. *International Journal of Science and Mathematics Education*, *13(6)*, 1211 1233.

Plöger, W., Scholl, D., & Seifert, A. (2015). Analysekompetenz – ein zweidimensionales Konstrukt?! Unterrichtswissenschaft [Analytical competence – a two-dimensional construct?!]. *Zeitschrift Für Lernforschung*, *43(2)*, 166–184.

Riese, J., & Reinhold, P. (2012). Die professionelle Kompetenz angehender Physiklehrkräfte in verschiedenen Ausbildungsformen [Professional competence of student teachers enrolled in different teacher training programs]. *Zeitschrift für Erziehungswissenschaft*, *15(1)*, 111–143.

Riese, J., Kulgemeyer, C., Zander, S., Borowski, A., Fischer, H., Gramzow, Y., Reinhold, P., Schecker, H. & Tomczyszyn, E. (2015). Modellierung und Messung des Professionswissens in der Lehramtsausbildung Physik [Modelling and measurement of professional knowledge in academic physics teacher education]. In S. Blömeke, & O. Zlatkin-Troitschanskaia (Eds.). *Kompetenzen von Studierenden: 61. Beiheft der Zeitschrift für Pädagogik [Competences of students in higher education]* (pp. 55–79). Weinheim: Beltz.

Sadler, P. M., Sonnert, G., Coyle, H. P., Cook-Smith, N., & Miller, J. L. (2013). The influence of teachers' knowledge on student learning in middle school physical science classrooms. *American Educational Research Journal*, 50(5), 1020 – 1049.

Seifert, A., & Schaper, N. (2012). Die Entwicklung von bildungswissenschaftlichem Wissen: Theoretischer Rahmen, Testinstrument, Skalierung und Ergebnisse [The development of Pedagogical Knowledge: Theoretical Framework, Test Instrument, Scaling and Results]. In J. König & A. Seifert (Eds.*), Lehramtsstudierende erwerben pädagogisches Professionswissen – Ergebnisse der Längsschnittstudie LEK [Pre-service teachers acquire pedagogical professional knowledge – results of the longitudinal study LEK]* (pp. 183–214). Münster: Waxmann.

Shavelson, R. J., & Stern, P. (1981). Research on Teachers' Pedagogical Thoughts, Judgements, Decisions, and Behavior. *Review of Educational Research*, *51(4)*, 455–498.

Shulman, L. (1986). Those Who Understand: Knowledge Growth in Teaching. *Educational Researcher*, *15(2)*, 4–14.

Stender, A. (2014). *Unterrichtsplanung: Vom Wissen zum Handeln. Theoretische Entwicklung und empirische Überprüfung des Transformationsmodells der Unterrichtsplanung [Lesson planning: From knowledge to action. Theoretical development and empirical verification of the transformation model of lesson planning].* Berlin: Logos.

van Driel, J., Biejaard, D. & Verloop, N. (2001). Professional Development and Reform in Science Education – The Role of Teachers' Practical Knowledge. *Journal of Research in Science Teaching*, *38(2)*, 137–158.

Vogelsang, C., Bowoski, A., Buschhüter, D., Enkrott, P., Kempin, M., Kulgemeyer, C., Reinhold, P., Riese, J., Schecker, H. & Schröder, J. (2019). Entwicklung von Professionswissen und Unterrichtsperformanz im Lehramtsstudium Physik – Analysen zu valider Testwertinterpretation [Development of Professional Knowledge and teaching skills in Academic Physics Teacher Education – analyses regarding the valid interpretation of test scores]. *Zeitschrift für Pädagogik*, *65(4)*, 473–491.

Volmer, M., Pawelzik, J., Todorova, M. & Windt, A. (2019). Reflexionskompetenz von Sachunterrichtsstudierenden im Praxissemester. In C. Maurer (Ed.), *Naturwissenschaftliche Bildung als Grundlage für berufliche und gesellschaftliche Teilhabe [Scientific Education as a basis for vocational and social participation]* (pp. 321–324). University of Regensburg.

Woitkowski, D. & Borowski, A. (2017). Fachwissen im Lehramtsstudium Physik [Content Knowledge in Academic Physics Teacher Education]. In H. Fischler & E. Sumfleth (Eds.), *Professionelle Kompetenz von Lehrkräften der Chemie und Physik [Professional Competence of chemistry and physics teachers]* (pp. 57–76). Berlin: Logos.