Professional knowledge affects action-related skills: The development of preservice physics teachers' explaining skills during a field experience

Christoph Kulgemeyer1 | Andreas Borowski2 | David Buschhüter2 | Patrick Enkrott2 | Maren Kempin1 | Peter Reinhold3 | Josef Riese4 | Horst Schecker1 | Jan Schröder4 | Christoph Vogelsang5

1Institute of Science Education, Physics Education Department, University of Bremen, Bremen, Germany
2Institute of Physics and Astronomy, University of Potsdam, Potsdam, Germany
3Physics Education, Department of Physics, University of Paderborn, Paderborn, Germany
4Institute of Physics IA, Physics Education Department, RWTH Aachen University, Aachen, Germany
5Center for Educational Research and Teacher Education, University of Paderborn, Paderborn, Germany

Correspondence
Christoph Kulgemeyer, Institute of Science Education, Physics Education Department, University of Bremen, Otto-Hahn-Allee 1, 28359 Bremen, Germany.
Email: kulgemeyer@physik.uni-bremen.de

Abstract
Professional knowledge is an important source of science teachers' actions in the classroom (e.g., personal professional content knowledge [pedagogical content knowledge, PCK] is the source of enacted PCK in the refined consensus model [RCM] for PCK). However, the evidence for this claim is ambiguous at best. This study applied a cross-lagged panel design to examine the relationship between professional knowledge and actions in one particular instructional situation: explaining physics. Pre- and post a field experience (one semester), 47 preservice physics teachers from four different universities were tested for their content knowledge (CK), PCK, pedagogical knowledge (PK), and action-related skills in explaining physics. The study showed that joint professional knowledge (the weighted sum of CK, PCK, and PK scores) at the beginning of the field experience impacted the development of explaining skills during the field experience ($\beta = .38^{**}$). We interpret this as a particular
The effects of field experiences (also called “practicums” or “school internships”) on preservice science teachers’ professional development have been researched extensively but mostly in descriptive studies (Holtz & Gnambs, 2017). In recent years, the once-prevalent assumption for teacher education that field experience is the key to success in general, has been replaced by a more differentiated view (Cohen, Hoz, & Kaplan, 2013). However, field experiences still play a major role for one of the most important challenges of teacher education (and not just in science): the theory-practice gap (e.g., Cohen et al., 2013) or as Holtz and Gnambs (2017) put it: “How can abstract academic learning be transformed into applied instructional practice?” (p. 669). This question is at the core of the notion of science teachers’ professional knowledge. Professional knowledge is assumed to be “abstract academic” knowledge that teachers need to achieve high-quality “instructional practice” (Alonzo, Berry, & Nilsson, 2019; Kulgemeyer & Riese, 2018). Zeichner, Payne, and Brayko (2012) see the role of field experiences exactly as moderators for the theory-practice gap: they are supposed to foster the transformation of knowledge achieved at the university into the skills necessary for the profession of a teacher in schools. From an empirical point of view, it remains an open question whether (a) professional knowledge is a source (disposition) that affects science teachers’ action-related skills important for instructional quality and (b) field experiences help to transform science teachers’ professional knowledge into these action-related skills.

In this article, we contribute to both of these questions by applying a cross-lagged panel analysis of the development of knowledge and instructional quality during a preservice physics teachers’ one-semester field experience. The focus lies on question (a), the contribution to question (b) remains more descriptive. In addition, we focus on one particular instructional situation: explaining physics. The reason for choosing this specific instructional situation is that much research has been conducted on instructional explanations in physics teaching recently so that concepts and assessment instruments are at hand. In particular, we can make use of performance assessments that offer a higher validity than common approaches to assess action-related skills through self-reports.
For the study reported in this article, professional knowledge is understood as the knowledge achieved in academic science teacher education (cf. personal pedagogical content knowledge (PCK; Alonzo et al., 2019)). It comprises CK, PCK, and pedagogical knowledge (PK; cf. Shulman, 1987) gained through university studies. Action-related skills (and explaining skills in particular) are those skills a teacher possesses that show up in classroom action and can be observed directly in terms of instructional quality. The theoretical background of this notion will be discussed below based on the refined consensus model (RCM) of professional CK (PCK; Hume, Cooper, & Borowski, 2019). The present paper will provide evidence for the assumption that personal PCK affects enacted PCK in explaining situations.

2 | THEORETICAL BACKGROUND

2.1 | The effects of field experiences on the development of action-related skills

Field experiences are an essential part of academic teacher education in various countries. Holtz and Gnambs (2017) point out that the purpose of these field experiences often is bridging the theory-practice gap between professional knowledge achieved at universities and teachers’ actions in the classroom. According to Blömeke, Gustafsson, and Shavelsson's (2015) model of competence (MoC), preservice teachers might develop the skills they need to successfully master practical situations in their profession during field experience, and even more, their professional knowledge helps them with this challenge. However, this is certainly not an easy task to accomplish. The naïve assumption that “more field experience leads to a better application of theoretical knowledge” has been questioned many times (e.g., Tabachnick & Zeichner, 1984), but most of the studies concerned with the effects of field experiences in teacher education have been merely descriptive (Wilson & Floden, 2003). Descriptive studies, however, contribute little to the goal of reliably identifying factors that make field experiences successful. Cohen et al. (2013) reviewed 113 studies concerned with field experiences in teacher education and found generally positive effects, for example, regarding self-confidence, observation skills, and action-related skills. In general, preservice teachers seem to think of themselves as more competent after a field experience (Besa & Büdcher, 2014). However, as Holtz and Gnambs (2017) point out, most of these studies rely on self-reports, which questions their validity in many ways. For example, self-reports often lead to people overestimating their skills (Oeberst, Haberstroh, & Gnambs, 2015). Holtz and Gnambs (2017), therefore, used a rating system with different perspectives that included self-reports as well as ratings from the pupils and the experienced teachers mentoring the interns. Overall, they found that the interns’ instructional quality increased during the semester-long field experience. Still, their study lacks a proximal measurement of action-related skills. It relies on more or less high-inferential ratings, which are an indirect measurement of instructional quality. The study reported in this article, therefore, aims at a more proximal measurement to increase the validity of the outcomes.

2.2 | Professional knowledge as a source (disposition) of science teachers’ actions

Over the last decades, science teachers’ professional knowledge has been researched extensively in science education (e.g., Abell, 2007; Fischer, Borowski, & Tepner, 2012; Fischer, Neumann,
Labude, & Viiri, 2014; Hume et al., 2019; Peterson, Carpenter, & Fennema, 1989; Van Driel, Verloop, & De Vos, 1998). Most of this research is based on Shulman’s (1987) fundamental considerations as a starting point and relies on three areas of professional knowledge in particular: CK, pedagogical CK (PCK), and PK (e.g., Baumert et al., 2010; Cauet, Liepertz, Kirschner, Borowski, & Fischer, 2015; Hill, Rowan, & Ball, 2005). Briefly, CK is assumed to be the factual knowledge of the content selected for teaching. PCK is described as the knowledge of how to teach this particular content, and PK comprises more general knowledge of how to act in instructional situations. PCK, in particular, has been discussed lately, resulting in the RCM of PCK (Hume et al., 2019) and in a notion of PCK that regards CK and PK as “foundational to teacher PCK in science” (Carlson & Daehler, 2019, p. 82).

Professional knowledge as a whole—comprising both declarative and procedural knowledge—has been thought of as a disposition for teachers’ action-related skills and, therefore, as crucial for achieving a high instructional quality that finally results in student achievement (Diez, 2010; Terhart, 2012). However, it is still unclear whether CK, PCK, and PK are major factors that influence teachers’ actions (Cauet et al., 2015; Vogelsang, 2014). Nevertheless, many teacher education systems focus on developing knowledge in these three areas. The German teacher education system, in particular, does so by offering distinct courses for all three areas from the beginning of academic studies (e.g., courses in physics, physics education, and pedagogy). A lack of evidence regarding the impact of the three areas on teachers’ action-related skills and instructional quality, therefore, would be a missing argument for the core elements of academic teacher education.

Indeed, the evidence for this relationship between professional knowledge and instructional quality is ambiguous at best. One prominent example is the study conducted by Hill et al. (2005). They successfully used a joint knowledge of CK and PCK (“CK for teaching mathematics”) to predict both instructional quality and student learning (Hill et al., 2008). However, Delaney (2012) used the same test instruments that Hill et al. (2005) used but did not find a significant correlation. In mathematics, the Professional Competence of Teachers, Cognitively Activating Instruction, and Development of Students’ Mathematical Literacy study showed that student achievement benefits from teachers’ PCK, which is mediated by cognitive activation (a major factor of instructional quality). CK also has a positive influence mediated by PCK (Baumert et al., 2010). The Professional Knowledge in Science (ProwiN) project did not find a relationship between teachers’ CK or PCK and instructional quality (operationalized by cognitive activation) or student achievement (Cauet et al., 2015). In addition, they did not find a relationship between PCK and CK on the one hand and the interconnectedness of the content structure on the other hand—while the interconnectedness affected students’ achievement (Liepertz & Borowski, 2018). Similar results were achieved by the Quality of Instruction in Physics (QuIP) project (Fischer et al., 2014; Keller, Neumann, & Fischer, 2016). Ergönenc, Neumann, and Fischer (2014) reported a low but positive correlation between PCK and learning outcome \(r = .270; \ p < .05\). Students had the greatest learning gain compared to their Swiss and German peers, but the Finnish teachers had a lower PCK than their German and Swiss colleagues.

### 2.3 The RCM of PCK and the MoC

Recently, two different theoretical approaches have been developed to understand the relationship between professional knowledge and action-related skills. The first approach is rather new
and related to the RCM of PCK in science education (Hume et al., 2019). The RCM was developed by an international group of two dozen science education researchers to find a consensual meaning of the construct PCK (Carlson & Daehler, 2019). As a result, it proposes different realms of PCK that are important to understand the relationship between professional knowledge and action-related skills: collective PCK (held by multiple specialized educators in general), personal PCK (the personal knowledge held by an individual educator), and enacted PCK (the knowledge and skills that only exist in the actions that are necessary to perform in teaching situations) (Carlson & Daehler, 2019).

CK and PK, in this sense, are regarded as knowledge bases that influence collective PCK (Figure 1). The collective PCK is an “amalgam of multiple science educators’ contributions” (Carlson & Daehler, 2019, p. 88). It influences personal PCK and vice versa. On the one hand, personal PCK is an individual subset of the collective PCK, while on the other hand, a teacher might decide to communicate his or her personal PCK in a way which contributes to the collective PCK. This “two-way knowledge exchange” (Carlson & Daehler, 2019, p. 82) is a characteristic of the whole model, which means that personal PCK impacts the enacted PCK and vice versa. The enacted PCK is defined as “the specific knowledge and skills utilized by an individual teacher in a particular setting, with a particular student or group of students, with a goal for those students to learn a particular concept […]” (Carlson & Daehler, 2019, p. 83). Therefore, it is the knowledge that is utilized in the practice of science teaching. Paper-and-pencil tests are not suitable tools to measure enacted PCK because it can only be observed in teachers’ actions. Paper-and-pencil tests are, however, common and suitable for measuring personal PCK. Alonzo et al. (2019) describe a possible mechanism to transform personal PCK into enacted PCK; the former is the main source that affects the latter, however, through reflection in or on action, the reverse applies. The question of whether professional knowledge achieved in academic teacher education affects teachers’ actions in the classroom—the core question to which this study wants to contribute—might be rephrased in terms of the RCM as follows: does personal PCK affect enacted PCK? In the sense of the RCM, not only the personal PCK affects the enacted PCK but also vice versa. Answering this question would contribute to evaluating a core concept of the RCM.

The second theoretical approach is based on the more general MoC by Blömeke et al. (2015) (Figure 2). Blömeke et al. (2015) developed their model intending to define a term widely used in educational measurement more accurately: competence. They understand competence as the “latent cognitive and affective-motivational underpinning of domain-specific performance in varying situations” (Blömeke et al., 2015, p. 3) and discuss the concept of competence in detail.

![FIGURE 1](image_url) Parts of the refined consensus model of pedagogical content knowledge (PCK) that have been used in the present paper. Content knowledge (CK), pedagogical knowledge (PK), curricular knowledge, assessment knowledge, and knowledge of students are foundational for PCK.
They argue that a dichotomy between competence (as a latent trait) and performance (as an observable behavior) overlooks how “knowledge, skills and affect are put together to arrive at performance” (Blömeke et al., 2015, p. 6). Competence, therefore, should be regarded as a continuum between disposition and observable performance. In essence, they regard cognitive resources (e.g., professional knowledge) and affective-motivational factors as dispositions that can be transformed into action-related skills (and refer to them as “situation-specific skills” necessary to master a particular situation) which finally result in performance. Performance is the observable outcome of situation-specific skills that consist of the perception of a situation, interpretation, and decision-making. The core question of this model is “[…] whether and how persons who possess all of the resources belonging to a competence construct are able to integrate them, such that the underlying competence emerges in performance” (Blömeke et al., 2015, p. 6).

The underlying idea is the same as the idea of the RCM; situation-specific skills are nearly identical to Alonzo et al.’s (2019) enacted PCK. However, in the sense of the MoC, a differentiation between CK, PK, and PCK makes sense. The RCM regards CK and PK as foundational for PCK while the MoC would support thatCK, PK, and PCK are at the same level, and all three areas impact action-related skills. Furthermore, in the sense of the RCM, it is easy to also imagine collective, personal, and enacted forms of CK and PK, based on how close to the actual classroom performance they are. Both models represent a continuum between theory (disposition/personal PCK) and practice (action-related skills or even performance/enacted PCK). The underlying idea of both models is that the theoretical knowledge positively influences the way teachers act in classroom situations. The main difference is that in the sense of the RCM the enacted PCK also influences the personal PCK while the MoC does not propose a concept of how the dispositions could be further developed.

The reasons for the mixed results regarding the relationship between professional knowledge and action-related skills or instructional quality have been discussed. Aufschnaiter and Blömeke (2010), for example, claim that validity issues in testing are part of the problem: paper-and-pencil tests might not be valid for measuring declarative knowledge and might not cover the probably more relevant procedural knowledge.

It is important to point out that despite the relationship between personal and enacted knowledge (in the sense of the RCM) or professional knowledge and action-related skills (in the sense of the MoC), all of the studies reported above conducted research on the validity of their test instruments and provided arguments, especially for their content validity. This means that these instruments reflect the content of teacher education as it is described in academic curricula. Indeed, they are adequate instruments to reflect on the (personal) knowledge that is a topic in science teacher education in the countries in which they conducted the studies. Of course, teacher education contains more than just professional knowledge (e.g., attitudes, beliefs, etc.; e.g., Weinert, 2001), but professional knowledge is a key element nonetheless. Does that suggest

**FIGURE 2** The model of competence by Blömeke et al. (2015) as it has been used in the present paper. Knowledge and affect-motivation are disposition, action-related skills are situation-specific skills, performance is the observable variable.
that this key element of teacher education does not influence teachers’ actions in the classroom? The other side of the measurement turned out to be an even bigger problem: it is challenging to measure instructional quality and action-related skills appropriately (Kulgemeyer & Riese, 2018). (cf. section 3.5)

2.4 Explaining physics and instructional quality

Explaining has often been described as a core element of teaching as it is a basic skill that enables communication with students (Geelan, 2012; Kulgemeyer, 2018; Wittwer & Renkl, 2008). Explaining with the goal of “engendering comprehension” (Gage, 1968) is the communicative process of proposing explanations and adapting them according to the students’ prerequisites and their comprehension (Kulgemeyer & Schecker, 2013). It is important to highlight that “explaining” in this sense should not be confused with transmissive teaching or didactical lecturing. It is a constructivist interaction between teachers and students, and even a well-crafted explanation can only increase the probability that students construct meaning from the displayed information.

The process of explaining needs to be distinguished from the more prominent notion of the term “explanation” concerning the nature of science and the well-researched area of argumentation in science classroom.

There have been discussions about the commonalities and differences between explanation and argumentation (Berland & McNeill, 2012; Osborne & Patterson, 2011). Both are similar concerning their structure and their epistemic nature. For example, the “covering law model” of explanation (Hempel & Oppenheim, 1948) shares commonalities with Toulmin’s (1958) argumentation patterns. However, explaining in the classroom is a communicative process that requires skills other than constructing a scientific explanation, which often relies on a logical connection between a phenomenon and an underlying law (even though numerous other models of explanation have been proposed). Treagust and Harrison (1999) distinguish between scientific explanations (which might follow the covering law model) and science teaching explanations as a communicative attempt to enable students to construct knowledge from teachers’ (verbal) actions. In this sense, explaining needs to adapt to the students’ prior knowledge, beliefs, and interests (Wittwer & Renkl, 2008). Kulgemeyer and Schecker (2013) highlight that a good argumentative structure might lead to acceptance or conviction but not always to understanding. Understanding, however, is the goal of an instructional explanation (Wittwer & Renkl, 2008).

Kulgemeyer (2018) presents a systematic literature review of factors that influence the effectiveness of an instructional explanation in the classroom. The communicative process of explaining has been researched in science education (e.g., Geelan, 2012; Sevian & Gonsalves, 2008) and educational psychology (e.g., Roelle, Berthold, & Renkl, 2014; Wittwer & Renkl, 2008). Wittwer and Renkl (2008) identify adaptation to the explainees’ knowledge as the crucial factor in effective explaining. Kulgemeyer and Tomczyszyn (2015) proposed a model of explaining in science teaching that has been researched for its validity (Kulgemeyer & Schecker, 2012; Kulgemeyer & Schecker, 2013; Kulgemeyer & Tomczyszyn, 2015). This model operationalizes how adaptation to the explainee might work in a science classroom and proposes four tools for adaption (representation forms, context/examples, mathematization, and code).

The process of explaining is modeled in a constructivist sense as follows: (Kulgemeyer & Schecker, 2009). If an explainer wants to communicate science content, two main points need
to be considered: (a) “what is to be explained?” (perspective: the structure of the science matter) and (b) “who is it to be explained to?” (perspective: adaptation to the prerequisites of the explainee). The explainer makes assumptions about both points based on professional knowledge (e.g., PCK about misconceptions). However, the explainer needs to evaluate these assumptions, for example, by asking appropriate questions. The explainee is free to accept or reject the explainer’s attempt to explain. The explainee, therefore, plays an active role in the explaining process and gives signs of understanding or misunderstanding, both verbally and nonverbally. The explainer must perceive and interpret those signs and adapt the communicative attempts accordingly. The four tools for adaptation are: (a) the level of mathematization, (b) representational forms (e.g., diagrams or realistic pictures), (c) contexts and examples (e.g., everyday contexts or physics textbook examples), and (d) the code of the (verbal) language used (e.g., informal language, technical language, or language of education).

Following this model, the instructional quality of explaining situations in teaching can be measured by rating how well an explainer diagnoses understanding or misunderstanding and how well he or she uses the four tools for adaption in regard to the explainees’ needs.

2.5 | A framework for instructional quality in explaining situations

Consequently, (a) adaptation to prior knowledge and (b) the four tools for adaption have emerged from a systematic review of empirical literature from science education and educational psychology as two out of seven core ideas of effective instructional explanation (Kulgemeyer, 2018). In this study, we use the seven core ideas from the framework for instructional quality in science explaining situations. The other five core ideas are:

Highlight the relevant parts: It seems to be important to clearly communicate that the explained content is relevant to the explainees. Studies show that this results in engaging explainees in a deeper processing of the information. The literature describes (a) prompts (Berthold, Röder, Knörzer, Kessler, & Renkl, 2011; Chi, Siler, Jeong, Yamauchi, & Hausmann, 2001); (b) responsive explaining (e.g., Diakidoy, Kendeou, & Ioannides, 2003); and (c) remedial explanations (Roelle et al., 2014; Sánchez, García Rodicio, & Acuña, 2009) as possible ways to highlight relevancy.

1 Structure the explaining: A “rule-example structure” is probably more adequate to achieve CK (Tomlinson & Hunt, 1971) while an “example-rule structure” is probably more fitting to develop action-related skills (Champagne, Klopfer, & Gunstone, 1982).

2 Minimal explanations: Explanations should be limited to what matters (Black, Carroll, & McGuigan, 1987). They should be free from redundancies, which suggest low relevance (Anderson, Corbett, Koedinger, & Pelletier, 1995). In addition, they should aim for a low cognitive load and a high coherence (Wittwer & Ihme, 2014; Kulgemeyer & Starauschek, 2014).

3 Starting condition: Teacher explanations have a place in the science classroom if students’ self-explanations are likely to fail (Renkl, 2002). This is the case for complex principles that are new to students or if students hold many misconceptions about the principle. Students tend to stop self-explaining activities when they think they have reached an understanding; however, that is sometimes deceptive and might result in even deeper misconceptions. Teacher explanations might help to avoid this so-called “illusion of understanding” (Chi, Bassok, Lewis, Reimann, & Glaser, 1989; Chi, de Leeuw, Chiu, & LaVancher, 1994).
4 Diagnose understanding: As described above, diagnosing understanding is crucial for further adaption. An instructional explanation, therefore, should always be followed by a diagnosis of understanding that allows further adaption. Students also need time to use the explained information to solve related problems in well-crafted learning tasks (Webb, Ing, Kersting, & Nemer, 2006; Webb, Troper, & Fall, 1995).

3 | METHODS

3.1 | Research question

The core research question of the present study is: Does professional knowledge (personal PCK) affect the development of explaining skills (enacted PCK in explaining situations)?

**Hypothesis 1** Professional knowledge (personal PCK) affects the development of explaining skills (enacted PCK in explaining situations) during a one-semester field experience. Professional knowledge, therefore, is a source (disposition) for these particular action-related skills.

**Hypothesis 2** Explaining skills (enacted PCK in explaining situations) impact the development of professional knowledge (personal PCK). The action-related skills of explaining physics, therefore, are a source for personal professional knowledge.

We want to highlight that the research question focuses on the relationship between professional knowledge and the development of explaining skills in a field experience. This is a very particular form of relationship between the two variables and requires an explanation. From a theoretical perspective, it makes sense to focus on skill development during a field experience because the transformation of professional knowledge into practical skills is exactly the bridge for the theory-practice gap that field experiences are supposed to offer (e.g., Holtz & Gnambs, 2017). If professional knowledge were transformed into explaining skills during the field experience, we would observe an impact of the professional knowledge at the beginning of the field experience on the development of explaining skills (path e in Figure 3) (see also “Study Design” below).

Besides, Hypothesis 1 is based on both theoretical approaches discussed above and, in particular, the RCM. Blömeke et al. (2015) regard professional knowledge as the source (disposition) for action-related skills. Alonzo et al.’s (2019) understanding of enacted PCK supports the hypothesis as well, as they assume personal PCK (pPCK) to be the source of enacted PCK (ePCK): “[...] these intuitive actions (planning, teaching, and reflecting) are all influenced by a science teacher’s pPCK; however, at the moment, they exist as ePCK” (p. 274). The authors also assume that ePCK might influence pPCK by reflecting about situations: “A conscious process may transform ePCK into pPCK in a form that can be articulated by the teacher. This transformation happens primarily through reflection in, or on, a science teaching episode as intuition and experiences become part of future knowledge that can be explicitly drawn upon in planning, teaching, and reflection” (p. 275).

This leads to Hypothesis 2: by reflecting on the reasons for the success or failure of the instructional activities, preservice teachers might develop their personal professional
knowledge from their enacted professional knowledge during a field experience. In this case, we would observe a relationship between explaining skills at the beginning of the field experience and the development of professional knowledge during the field experience (path f in Figure 3).

Therefore, we will investigate the research question by testing both hypotheses: does professional knowledge affect the development of action quality—or does action quality result in higher professional knowledge? A cross-lagged panel design (a design with repeated measurement of both possible source and skill variables) allows both hypotheses to be tested at the same time by researching four criteria (see “Study Design” below).

3.2 | Study context and design

3.2.1 | Study context

This study took place in the context of the German teacher education system. In Germany, most federal states implement a so-called “practical semester” in their study programs, a one-semester (approximately 5 months) field experience at schools under the supervision of a university lecturer and the mentorship of experienced teachers. During this field experience, the preservice teachers mainly work at the schools but still have some courses (usually 1 day per week) at the university to support them in their practice. The implementation of this idea differs not only from university to university but also from school to school in which the students work. Hoffman et al. (2015) point out that the role of the mentors also differs largely, for example, influenced by their national or local context (like the school district). They are mostly underprepared for their role and experience a tension between mentoring preservice teachers on the one hand and being responsible for their students on the other. For this study, we, therefore, decided to administer a questionnaire to the preservice teachers to learn more about how the field experience worked out for them. The basic structure, however, is given by the university curricula: they all spend most days of the week in a school, are assigned experienced physics teachers as mentors, attend supporting seminars at their universities in science education, teach physics, and are observed by their university educators at least once to receive feedback on their teaching. Both the support of the experienced teachers and the supervising university lecturers are of disciplinary nature—in case of the present paper physics education. In this setting, we regard it as highly likely that communicative situations with students will appear to allow the preservice teachers to develop their explaining skills. Below, we will describe how the field experience worked.
3.2.2 Study design: Cross-lagged panels

Cross-lagged panel designs are research designs that help to investigate relationships between variables in detail. They are based on longitudinal studies with the repeated measurement of the same set of variables (“waves”).

The study design follows a cross-lagged panel design with two waves (Figure 3). Before and after the field experience, the students were tested (see “instruments” below) for their CK, PCK, and PK (professional knowledge) on the one hand and their explaining skills on the other hand. Additionally, before the field experience, they filled out a questionnaire on demographics (e.g., teaching experience, etc.). After the field experience, they completed another questionnaire about their experiences in their schools. The testing time was approximately 3 hr and 30 min before and after the field experience.

Reinder (2006) review of the literature on cross-lagged panels claims that such a design enables three criteria to be observed that suggest a causal relationship between variables. These criteria might not be sufficient to claim a causal relationship between professional knowledge and action-related skills as long as the mechanism how knowledge is transformed into action remains unclear. However, these three criteria still support the claim that professional knowledge is necessary for action-related skills in general. We, therefore, use the three criteria to judge whether or not professional knowledge is a disposition for action-related skills:

1 The cause should happen before the effect. This means that the measurement of the cause (in this case, professional knowledge) should happen before the measurement of the effect (the development of explaining skills) to research this criterion. This criterion is applicable to the research design because, obviously, the pretest happens before the posttest. Both hypotheses can be tested: professional knowledge (pretest) is measured before the explaining skills (posttest) (Hypothesis 1) but also vice versa (Hypothesis 2). It is important to highlight which cause and which effect concerns our research design: we research whether or not the cause “professional knowledge” affects the effect “development of explaining skills.” This is a particular aspect of a relationship between the cause “professional knowledge” and the effect “explaining skills” (see discussion above). Explaining skills, therefore, are assumed as skills that exist apart from professional knowledge but are developed from professional knowledge in practice.

2 Cause and effect should be correlated. This criterion is fulfilled if the prior state of professional knowledge explains the development of explaining skills during the field experience (Hypothesis 1) (path e) or vice versa (Hypothesis 2, path f).

3 The cause should be the only or the main factor explaining the effect. This criterion is often regarded as fulfilled if it is theoretically sound that the most important variables are included in the measurement and the correlation between professional knowledge and the increase in explaining skills is significantly higher than the correlation between explaining skills and the increase in professional knowledge (Hypothesis 1, path e > path f) or vice versa (Hypothesis 2, path f > path e).

The fourth criterion for possible causal relationships comes from Granger (1969). We also use it to decide whether or not professional knowledge is necessary for action-related skills.

An effect should be higher correlated with the cause than with its prior value. This means that the future value of a variable is better predicted with the assumed cause than with the prior value of the effect variable. This criterion can be regarded as fulfilled if the correlation between
professional knowledge and the increase in explaining skills is higher than the autocorrelation between the pretest and posttest (Hypothesis 1, path e > path c) or vice versa (Hypothesis 2, path f > path d).

One of the strengths of this research design is that it is possible to compare both hypotheses of the research question at the same time. It also allows us to describe the effects of the intervention based on pre- and posttests. However, due to a missing control group, that will remain descriptive. The analysis of the data from a cross-lagged panel is usually based on partial correlations controlling the pretests or structural equation modeling. We will discuss the methods for data analysis below.

3.3 | Sample

The sample consisted of $N = 47$ (pre and post: $N = 89$ cases, missing data from five cases) preservice teachers from four German universities (male: 20, female: 23, missing data from four preservice teachers). All of these preservice teachers held a bachelor's degree in physics education after completing a six-semester bachelor's program and were, on average, in their eighth semester ($M = 8.4; SD: 2.92$, ranging from 0 to 20). The field-experience was part of their master's program in physics education. Thirty-one of the preservice teachers had focused on physics as a subject in their “Abitur” (German high school degree, missing data from four students). Their second teaching subjects (German preservice teachers study two subjects) were mostly math ($N = 25$) or chemistry ($N = 7$). Ten of them had already participated in a field experience (ranging from 4 to 15 weeks).

The rather small sample size is due to the low number of students in German universities aiming to complete a degree in teaching physics. The sample, therefore, represents an actual census of the participants in the field experience at four German universities. This sample size is still appropriate as we expected to observe large effects due to the long duration of a 5-month intervention.

3.4 | Description of the intervention

The preservice teachers taught an average of 18.9 (ranging from 5 to 50) physics lessons during the field-experience, 3.7 (ranging from 0 to 20) of them in the area of mechanics, which was the area of physics on which the instruments focused. Additionally, they observed 64.4 (ranging from 12 to 140) physics lessons taught by experienced teachers during the field experience. They claimed to have prepared 4.6 hr (ranging from 1 to 15) for each lesson they taught and to have spent between 5 and 10 min (ranging from 0 to 30) before each lesson in talking with their mentor teacher about the preparation. After each lesson, they claimed to have spent between 5 and 10 min (ranging from 0 to 20) with their mentor teacher in reflecting on the lessons, mostly regarding classroom management and the tasks they gave their students. Teacher educators from the universities observed an average of 2.0 (ranging from 1 to 11) of their lessons and spent, on average, between 15 and 30 min (ranging from 0 to 90 min) in reflecting with them on those lessons afterward. The preservice teachers claimed to have already taught 9.2 physics lessons (ranging from 0 to 80) before the field experience. The preservice teachers, therefore, had gained practical experience, which was a prerequisite for this study.
3.5 | Measuring instructional quality and action-related skills

Traditionally, lesson videography has been used to measure instructional quality. For example, the QuIP project videotaped 69 physics teachers in a total of 92 lessons to examine the relationships between the teachers’ professional knowledge and instructional quality. The ProwiN study in physics videotaped a total of 70 lessons from 35 teachers. Keeping in mind that teachers’ professional knowledge is just one of many variables influencing instructional quality (cf. Helmke, 2006 or Seidel and Shavelson (2007)), it becomes obvious that with so many variables and so little lessons only a very large effect of professional knowledge on the instructional quality can be identified reliably (Kulgemeyer & Riese, 2018). As far as the two mentioned projects go, large effects were not found. The time and effort it takes to measure instructional quality might cause a systematic problem with the statistical power such a study can achieve.

To cope with these problems, Kulgemeyer and Riese (2018) used performance assessments to measure teachers’ practical skills. Performance assessments rely on a direct measurement of action, but this is under standardized conditions, often involving “standardized students” who are trained to play a specific role. These assessments, thus, are more proximal measurements of action-related skills. Although the test situation is standardized, a performance assessment situation has to be authentic for real teaching. In the sense of the RCM, performance assessments measure enacted knowledge. Alonzo et al. (2019) explicitly mention studies using performance assessments as examples for studies on enacted PCK (e.g., Kulgemeyer & Schecker, 2013). According to Blömeke et al.’s (2015) MoC, these assessments measure action-related (situation-specific) skills. To measure explaining skills comparably between different test persons, Kulgemeyer and Riese (2018) developed standardized simulated explaining situations where the test persons met students who had been trained to ask standardized questions about given topics. They found that PCK mediated the influence of CK on explaining skills and that both aspects of professional knowledge are sources (dispositions) of teachers’ practical skills. However, their research design had two limitations: (a) the data from this cross-sectional study could only identify correlations, and (b) the focus on explaining situations covers just one of many important classroom situations relevant for instructional quality. Regarding the second limitation, Kulgemeyer and Riese (2018) argue that constraining performance assessment to a specific situation allows for detailed analyses. The main advantage of their approach lies in the standardization of the situation. Many variables confounding the relationship between professional knowledge and instructional quality can be controlled; in this case, it is the method and content of students’ questions when they ask for explanations. Simulating explaining situations should be regarded as a starting point for more performance assessments involving different teaching situations (e.g., reflecting on physics lessons, conducting an experiment, etc.). Performance assessment is a promising tool because it allows deep insights into teaching situations and the controlled stimulation of teachers’ actions in these situations. Last but not least, performance assessments in controlled settings allow data to be gained from a larger number of test persons compared to video studies on whole lessons. Thus, performance assessments can even be used to analyze developments pre and post an intervention (as in the present paper).

3.6 | Instruments

We used established instruments in this study that have previously been used in other research projects (e.g., Kulgemeyer & Riese, 2018). The development of the instruments is described in
detail in Riese et al. (2015) and Vogelsang et al. (2019). The instruments for PCK, CK, and explaining skills all measure within the area of mechanics because we regard it as likely that this common area in German high school curricula is part of the preservice teachers’ duties during their field experience. The pretest and the posttests were identical with the exception that the problems they had to explain were different. Kulgemeyer and Tomczyszyn (2015) showed that the chosen topic has no significant effect on the test score.

3.6.1 | (Personal) PCK

This instrument was based on the PCK model by Gramzow, Riese, and Reinhold (2013). The model itself is based on other concepts that were also the basis of the RCM, for example, concepts from Lee and Luft (2008), Magnusson, Krajcik, and Borko (1999), and Park and Oliver (2008). The model, however, has been adapted to the German situation by relying on an additional curricular analysis of German academic teacher education concerning PCK. The focus was on PCK that students can acquire in German academic teacher education.

The paper-and-pencil test contained 43 items (see Figure 4 for a sample item), including both open situational judgment items and multiple-choice items. These 43 items were selected concerning the statistical fit and content validity. The test required 60 min in its final version. The test was analyzed for content validity (expert ratings), construct validity (using a dimensional analysis by applying the Rasch model and the analysis of a nomological network), and cognitive validity (using a think-aloud study). It was found to be reliable overall (EAP/PV reliability: 0.84) and reaches an appropriate interrater reliability ($\kappa = .87$). It consists of four reliable scales: (a) instructional strategies (EAP/PV reliability: 0.62), (b) students’ misconceptions and how to deal with them (EAP/PV reliability: 0.69), (c) experiments and teaching of an adequate understanding of science (EAP/PV reliability: 0.74), and (d) PCK-related theoretical concepts (e.g., conceptual change [Scott, Asoko, & Driver, 1992]) (EAP/PV reliability: 0.76).

3.6.2 | Pedagogical knowledge

To assess PK, we relied on an adapted short scale of an established instrument developed by Seifert and Schaper (2012). Seifert and Schaper (2012) also provide arguments for construct validity and content validity. Both the complete instrument and the short scale have been used in several studies (e.g., Mertens & Gräsel, 2018; Riese & Reinhold, 2012). The short-scale

![Figure 4](image-url)
addresses two aspects of PK: general instructional strategies and classroom management. Test time was 15 min, and the instrument is reliable ($\alpha = .76$).

An (open) sample item would be: *Name in note form at least five characteristics of good instruction.*

### 3.6.3 Content knowledge

The model for physics teachers’ CK consists of three scales that form reliable subscales of the test instrument: (a) school knowledge (physics knowledge from school textbooks) (EAP/PV reliability: 0.84), (b) deeper school knowledge, a deeper understanding of physics knowledge from a school textbook (e.g., knowledge of different procedures of solution) (EAP/PV reliability: 0.76), and (c) university knowledge (physics knowledge from a university textbook) (EAP/PV reliability: 0.76). The test consisted of 48 multiple-choice questions and took 60 min. Content validity has been researched based on a curricular analysis of the participating universities and common textbooks. Similar to the instrument for PCK, construct validity was tested using Rasch analysis. Cognitive validity was researched based on a think-aloud study. The instrument turned out to be reliable in our study (EAP/PV reliability: 0.84). A sample item is given in Figure 5.

### 3.6.4 Explaining skills

We assessed explaining skills using a real teaching situation simulation (a performance assessment of explaining skills [Kulgemeyer & Schecker, 2013; Kulgemeyer & Riese, 2018]). In such a situation, the preservice teachers explain physics phenomena to a single high school student, either “Why does a car skid out of a sharp turn on a wet road?” (physics background: friction, forces, circular motion); or “Why could it theoretically work to blow up an asteroid approaching the earth into two halves to save the planet?” (physics background: conservation of momentum, superposition principle). Dialogic situations are a common part of science teaching, for example, when students are working in groups or on their own and teachers help them. It is a basic science communication situation with one expert and one novice; still, it is somewhat different from when a teacher explains in front of a whole class.

Forces act upon a cube of homogenous density as shown below. In one of the following cases it is not possible to describe the cube’s motion with the model of mass points. Which one?

**FIGURE 5** Sample item from the content knowledge (CK) test
The setting consisted of a 10-min preparation phase and 10 min of explaining. During the preparation phase, one of the two topics mentioned above was given randomly to the preservice teachers, as well as the task to explain this topic to a 10th-grade student, using the same set of materials, for example, diagrams of a well-defined complexity. The test persons were free to use any parts of the materials they wanted to use. Kulgemeyer and Tomczyszyn (2015) showed that neither the explaine nor the chosen topic affect the test scores.

During the explaining phase, the preservice teachers explained the topic to the students. They were filmed during this phase. The setting was standardized: a room with a table and chairs for the preservice teachers and the students, paper and pencils for taking notes and sketches, and a whiteboard. The most important standardization was that the students were trained to behave in a standardized way during the explaining situation. They gave standardized prompts (e.g., “Could you explain it again? It was a bit confusing for me.”).

We analyzed the resulting videos using an established instrument developed by Kulgemeyer and Riese (2018) based on the explaining model and the criteria for good explaining (perspective: adaptation to the prerequisites of the explaine). We also coded the scientific completeness and correctness of the explanation attempts (perspective: the structure of the science matter from the model of explaining). The final coding manual, drawn from Kulgemeyer and Riese (2018), consists of 24 categories (samples in Table 1). Explaining skills were rated using the sum of the categories occurring in a video.

This approach has been researched for validity in prior studies (Kulgemeyer & Schecker, 2013; Kulgemeyer & Tomczyszyn, 2015), e.g., the resulting measure predicts an experts’ decision on the better explaining quality between two videos (Cohen’s $\kappa = .78$). It also correlates with the rating of the trained students on the explaining quality (Pearson’s $r = .53; p < .001$). Besides, for construct validity, a nomological network was analyzed, and interview studies were conducted to ensure the setting’s authenticity. Two raters reached accordance, ranging between 73 and 100% for each category and were able to come to a consensus after a short communication. The method is reliable ($\alpha = .70$).

### 3.7 Methods of data analysis

To analyze the data of the cross-lagged panel, we relied on both partial correlations and a structural equation model (SEM). SEMs have become a common tool for analyzing cross-lagged panels (Reinder, 2006). The coefficients estimated between the measured variables represent standardized regression coefficients. To estimate coefficients, we used robust maximum-likelihood estimation to deal with missing data and, due to the low sample size, manifest values. Manifest models usually underestimate the relationship between traits, so this approach still allows us to identify relationships. For model fits, we present comparative fit index (CFI; very good fit CFI > .95) and root mean square error of approximation (RMSEA; required RMSEA < .05); however, based on the model structure and the low number of variables, the model fit is unlikely to fail these criteria.

The optimal sample size for an SEM has been discussed in the literature (cf. Wolf, Harrington, Clark, & Miller, 2013). Nunnally’s (1967) classic “rule of thumb” would require 10 cases per included variable, which would allow us to estimate a model with four variables. Therefore, we can only analyze the effects of a joint professional knowledge (two variables: pre and post) on explaining skills (two variables: pre and post) and not differentiate between the three areas of professional knowledge. This is a major limitation of our study, but still, our study is the first
one to analyze these effects in a cross-lagged panel. However, the analysis using structural
equation modeling should still be treated with care, and the analysis based on (partial) correlations is certainly the more careful choice.

In a nutshell, the sample size is limited but still large enough to observe the expected large effects for both the analysis of the cross-lagged panel and the development of knowledge and skills during the intervention based on $t$ tests.

The study variable “joint professional knowledge” is calculated as the weighted sum of all CK, PCK, and PK subscales ranging from 0 to 1. For this purpose, we took the test results from the subscales of CK, PCK, and PK (see above) and divided them by their maximum score. The sum of these scores was then divided by the maximum score of the joint scale to normalize the scale between 0 and 1. For the joint professional knowledge scale, therefore, each subscale has the same weight. This scale is required to analyze the cross-lagged panel (see the discussion about sample size above). It certainly is a straightforward approach to construct a joint scale of professional knowledge, but it has been done in other studies successfully in a similar way (Hill et al., 2005). We will discuss the decision on how to construct this score in the Discussion section below and present data for the subscales to show that the observed effects are not simply caused by the way the scale is constructed: if the effects on the joint scale can also be found in the subscales, we would interpret that as consistency in the evidence.

**TABLE 1** Examples of categories used to calculate measure of explaining skills (taken from Kulgemeyer and Riese (2018))

| Perspective from the model in Figure 1 | Category | Description |
|----------------------------------------|----------|-------------|
| Adaptation to the prerequisites of the explainee | Presenting concrete numbers in formulas | The explainer presents numbers as an example for applying the formula instead of keeping it abstract |
| Explaining physics concepts in everyday language | Explaining physics concepts in everyday language | The explainer avoids technical terms by describing the underlying concept with everyday terms: “Push” instead of “momentum”; “… and the water lifts it a bit” instead of “buoyancy” |
| Summary | Summary | The explainer summarizes the explanation |
| Encouragement | Encouragement | The explainer praises the explainee for good answers and encourages him/her to deal with the difficult parts of the explanation |
| Diagnosing understanding | Diagnosing understanding | The explainer diagnoses the success of the explanation by asking questions or giving tasks (NOT just: “Did you understand that?”) |

| Scientific correctness and completeness (examples from the sharp-turn-on-a-wet-road scenario) | Phenomenon named explicitly: The motorcycle corners | The explainer says explicitly that the motorcycle takes a sharp turn |
|----------------------------------------|-------------------------------------------------|-------------------------------------------------|
| | Physics reasoning of the sharp turn given | The explainer states that a force acts on the motorcycle |
| | The reasoning for the sharp turn is scientifically correct | The explainer states that the force points radially inward (centripetal force) |
Figure 6 shows the relationship between the study variables and the two models discussed above. We measure CK, PK, and PCK on the one hand and explaining skills on the other hand. CK, PCK, and PK constitute the joint professional knowledge scale we use for the analysis. The joint professional knowledge represents the disposition in the MoC and the personal PCK in the RCM. The explaining skills represent the action-related skills in the MoC and the enacted PCK in the RCM.

From the perspective of the MoC, a joint scale of professional knowledge is not a problem because all of its parts (the subscales) equally serve as dispositions for action-related skills. Even more, an impact from the joint professional knowledge on the action-related skills can be assumed but not vice versa.

From the perspective of the RCM, we would argue that such the joint professional knowledge scale reflects the fact that CK and PK are foundational to PCK and, therefore, can be integrated into one score. Besides, the subscales of our PCK instrument can also be interpreted as foundational for PCK in the sense of the RCM (e.g., knowledge of students is represented by the subscale students' misconceptions and how to deal with them).

A more cautious approach than the SEM is the analysis of correlations and, specifically, partial correlations controlling the data from the pretests (Reinder, 2006). Therefore, we also analyzed the cross-lagged panel data using Pearson's correlation coefficient (for relationships a, b, c, and d in Figure 3) and partial correlations controlling the pretests (for cross-lagged correlations e and f in Figure 3).

We will discuss the relationship by referring to the four criteria by Reinder (2006) and Granger (1969) presented above. To analyze the development of knowledge and skills during the intervention, we use t tests.

4 | RESULTS

4.1 | Descriptive statistics

As the first step, the study variables were analyzed regarding their mean score, standard deviation, and their range to identify possible cover or ground effects. This is especially important because of the long duration of the intervention. No such effects could be found (Table 2).
4.2 | Development of explaining skills during field experience

In the following, we report on an analysis of the intervention by using t tests. In general, neither the joint professional knowledge ($T(68) = −1.589 [p = .117]$) nor explaining skills ($T(87) = −1.15 [p = .252]$) changed significantly during the field experience.

A development for all preservice teachers regarding explaining skills, therefore, was not observed. However, if Hypothesis 1 is true, the level of professional knowledge at the beginning of the field experience determines whether or not a development of explaining skills occurs. A general development would not be observed if only those preservice teachers develop their explaining skills that hold an according level of professional knowledge before the field experience. To gain first evidence for this claim, we formed two groups of preservice teachers based on their professional knowledge score in the pretest: Group A with a high professional knowledge (50% of preservice teachers with the highest scores on the joint professional knowledge scale—higher than 0.51) and Group B with a low professional knowledge (50% with the lowest scores—lower than 0.51). Group A, indeed, showed a significant development in explaining skills during the field experience from an average score of 0.53 to 0.64 ($T(36.0) = −2.054 [p = .047]; d = 0.63$) while Group B showed no development at (0.54 (pre), 0.54 [post]) ($T(36) = −0.177 [p = .860]$).

4.3 | Does personal PCK affect the development of enacted PCK? Does enacted PCK affect the development of personal PCK?

To answer the main research Question 1, both Hypotheses 1 and 2 were tested using the four criteria described above. Table 3 presents manifest intercorrelations of the study variables, and Figure 7 shows the analysis of the cross-lagged panel data using both SEM and correlations.

Both approaches reveal similar relationships regarding the four criteria to determine an effect from professional knowledge on action-related skills in explaining situations:

| Table 2 | Descriptive statistics of study variables |
|---|---|---|---|---|
| | M | SD | Minimum | Maximum |
| CK1 | 0.59 | 0.123 | 0.33 | 0.87 |
| PCK1 | 0.50 | 0.121 | 0.17 | 0.71 |
| PK1 | 0.35 | 0.120 | 0.15 | 0.57 |
| Prof1 | 0.49 | 0.095 | 0.27 | 0.65 |
| Ex1 | 0.55 | 0.106 | 0.29 | 0.75 |
| CK2 | 0.63 | 0.123 | 0.40 | 0.84 |
| PCK2 | 0.56 | 0.125 | 0.29 | 0.85 |
| PK2 | 0.41 | 0.106 | 0.21 | 0.64 |
| Prof2 | 0.52 | 0.097 | 0.31 | 0.74 |
| Ex2 | 0.57 | 0.112 | 0.29 | 0.85 |

Abbreviations: 1, pretest; 2, posttest; CK, content knowledge; Ex, explaining skills; M, means; PCK, pedagogical CK, PK, pedagogical knowledge; Prof, joint professional knowledge (weighted sum of CK, PCK, and PK subscales); SD, standard deviation.
The cause should happen before the effect. This criterion is given by design for both Hypotheses 1 and 2; the possible reasons (professional knowledge/explaining skills) were measured before the possible causes.

Cause and effect should be correlated. For both the correlational analysis and the SEM, a relationship between professional knowledge before the intervention and the development of explaining skills can be identified ($r_{Prof1,Ex2} = .40^*; \beta = .38^{**}$).

The reason should be the only or the main factor explaining the cause. The correlation between professional knowledge (pre) and explaining skills (post) controlling explaining skills (pre) is higher than the correlation between explaining skills (pre) and professional knowledge (post) controlling professional knowledge (pre): $r_{Prof1,Ex2} = .40^* > r_{Ex1,Prof2\cdot Prof1} = -.06; \beta = .38^{**} > \beta = -.04$. This difference is significant ($Z = 1.69 [p = .045]$).

A cause should be higher correlated with the effect than with its own prior value. For the relationship between professional knowledge (pre) and the development of

**TABLE 3** Manifest intercorrelations of measured variables (Pearson’s $r$)

|       | CK1 | PCK1 | PK1 | Prof1 | Ex1 | CK2 | PCK2 | PK2 | Prof2 | Ex2 |
|-------|-----|------|-----|-------|-----|-----|------|-----|-------|-----|
| CK1   | —   |      |     |       |     |     |      |     |       |     |
| PCK1  | .43** | —   |     |       |     |     |      |     |       |     |
| PK1   | .26  | .56* | —   |       |     |     |      |     |       |     |
| Prof1 | .77** | .85** | .61** | —   |     |     |      |     |       |     |
| Ex1   | .12  | -.02 | -.11 | -.04  | —   |     |      |     |       |     |
| CK2   | .70** | .31  | .17  | .55** | .09 | —   |      |     |       |     |
| PCK2  | .43** | .74** | .49  | .66** | -.15 | .38* | —   |     |       |     |
| PK2   | .21  | .53** | .62** | .44  | -.00 | .12  | .58** | —   |       |     |
| Prof2 | .67** | .72** | .45* | .78** | -.05 | .72** | .91** | .55** | —   |     |
| Ex2   | .22  | .17  | .17  | .38*  | .23  | .17  | .13  | .11  | .17  | —   |

*Note: 1, pretest; 2, posttest; CK, content knowledge; Ex, explaining skills; PCK, pedagogical CK, PK, pedagogical knowledge; Prof, joint professional knowledge (weighted sum of CK, PCK, and PK subscales).

$p < .05$.

$p < .01$.

**FIGURE 7** Cross-lagged panel analysis. *$p < .05$, **$p < .01$. Left: Based on correlations (the cross-lagged correlation measures represent partial correlations controlling the pretests; the other measures are biserial correlations). Right: Based on a path model with standardized coefficients (robust maximum-likelihood full information-maximum-likelihood-estimation), comparative fit index (CFI) = 1.000, root mean square error of approximation (RMSEA) = 0.000.
explaining skills, this criterion is fulfilled:
\[ r_{\text{Prof1}, \text{Exp2} \cdot \text{Exp1}} = .40^* > r_{\text{Exp1}, \text{Exp2}} = .23; \beta = .38^{**} > \beta = .24. \]

To gain further evidence for the validity of our results, we used the fact that the PCK and the CK instruments provide reliable subscales that allow a detailed analysis. The joint scale for professional knowledge needs an empirical argument regarding validity. In Table 4, we present evidence for Criterion 3 described above. It contains partial correlations between the score in the respective subscale before the field experience and the explaining skills after the field experience controlling the explaining skills before the field experience (path e in Figure 3). This score is compared with path f in Figure 3—the correlation between the respective subscale of professional knowledge after the field experience and the explaining skills before the field experience controlling the particular subscale of professional knowledge before the field experience.

Table 4 shows that for all subscales with two exceptions (PCK-related theoretical concepts and deeper school knowledge) Criterion 3 has been fulfilled in tendency (\( r_{\text{SC1}, \text{Ex2} \cdot \text{Exp1}} > r_{\text{Exp1}, \text{Exp2}} \)).
However, just for the subscale “instructional strategies” from PCK and “school knowledge,” the partial correlation $r_{SC1,Ex1•Ex2}$ turns out to be significant.

Kulgemeyer and Riese (2018) highlight that not all aspects of professional knowledge are relevant for every situation of a teacher’s duty. In this case, one could assume based on the correlations that the subscales “PCK-related theoretical concepts” (from PCK) and “deeper school knowledge” (from CK) might not be essential to perform well in explaining situations. Thus, as a post hoc analysis, we recalculated the scale for professional knowledge and excluded these two subscales. The reanalysis of the cross-lagged panel using the “relevant professional knowledge” scale, as expected, led to stronger relations (Figure 8). All four criteria were fulfilled for this hypothetical scale as well.

5 | DISCUSSION

Regarding the research question, all four of the criteria support the hypothesis that professional knowledge affects the development of the skills of explaining physics during a field experience. We, therefore, assume that professional knowledge is an important source for the development of action-related skills in explaining physics during a field experience. In terms of the RCM, we would argue that pPCK affects ePCK but not necessarily vice versa (also keeping in mind Table 4 and the subscales of the PCK). The two-way knowledge exchange between pPCK and ePCK could not be observed but an impact of pPCK on ePCK has been found. Our results support this important assumption of the RCM.

In terms of the MoC, we would argue that PCK, CK, and PK are dispositions of the action-related (situation-specific) skill to explain physics; a teacher positively benefits in his or her performance in explaining situations from his or her PCK, CK, and PK. Our results support the MoC in this respect.

One major limitation of our study concerning the relationships within the RCM model proposed by Alonzo et al. (2019) is that these authors assume reflection on the action as a moderating factor that leads to the development of pPCK based on ePCK. The interconnected model of professional growth (Clarke & Hollingsworth, 2002) supports this assumption. This model regards reflection and enaction as the mediating processes that cause change in four possible domains: the personal and external domains, and the domains of practice and consequences. The pPCK might be considered as part of the personal domain and the ePCK as part of the domain of practice. Following this model, the process of how change occurs in pPCK would be reflection on ePCK, an external source of information, or salient outcomes. That highlights the importance of further research on how reflection might cause growth in pPCK; however, this is beyond the scope of this article. Indeed, we can identify whether or not the preservice teachers had the opportunity to reflect on their actions. We reported that they regularly reflected on their teaching with both their university educators and their mentor teachers (cf. Section 3.4) and we know how much time they spend on these attempts. However, the time they spend on reflection with their mentor teachers and their university educators did not turn out to be a useful predictor of the increase in the pPCK, then again, we do not know enough about the reflection process. First, we do not know how successful they were during their attempts to reflect; neither do we know if some simply reflected in detail on topics related to their PK while others reflected in detail on topics related to their PCK. More importantly, we do not know whether or not they reflected on ways to explain physics in particular as that information is not included in our data. This might have caused Hypothesis 2 to not be observed. Even though we did not
observe any impact of ePCK and pPCK we, therefore, would not recommend refining the RCM in this respect based on our study.

Two other major limitations of our study should be noted explicitly:

1. The sample size is rather small. This allows us to identify large effects, but the study should be repeated with a larger sample. Due to the high effort, a qualitative analysis of explaining skills requires ($N = 89$ videos were analyzed overall) and the fact that the sample already included nearly all students participating in the field experience from four universities, it was not possible to reach a higher sample size in this study.

2. We cannot be sure that we measured all the important variables. Other variables may explain more variance of explaining skills than professional knowledge does. Possible candidates are epistemological beliefs (e.g., Staub & Stern, 2002) and self-efficacy (e.g., Retelsdorf, Butler, Streblow, & Schiefele, 2010).

In general, we did not find a notable increase in explaining skills. This is a first sign that field experience might not cause the development of a particular action-related skill automatically. To our knowledge, our study is the first to reveal this effect using a proximal measurement of action-related skills. This result seems to contradict the assumption that action-related skills develop in field experiences, an assumption that is based on, among other things, self-reports (Holtz & Gnambs, 2017). Because of the direct measurement of action-related skills, we would claim a higher validity for our study. However, that contradiction is only true on a global level. Based on the analysis of the cross-lagged panel data, we could show that there actually is an increase in the action-related skill of explaining physics—but it depends on the level of professional knowledge before the field experience. We analyzed the relationship between the level of professional knowledge before the field experience and the action-related skills of explaining physics after the field experience controlling the level of action-related skills before the field experience. The correlational analysis and the SEM support the hypothesis that a higher professional knowledge at the beginning of the field experience leads to a better development of the action-related skill of explaining physics. In other words, preservice teachers with a certain level of professional knowledge will benefit from field experience. In our data, preservice teachers who started the field experience with a higher than-average professional knowledge showed a high development of their explaining skills ($d = 0.63$).

That result is supported by the analysis of the subscales in Table 4. The analysis suggests that for all subscales with two exceptions (PCK-related theoretical concepts, deeper school knowledge), Criterion 3 is fulfilled in tendency ($r_{SC1,Ex2} > r_{Ex1,SC2} > r_{Ex1,Ex2}$), and even for these two, it cannot be contradicted. However, just for the “instructional strategies” subscale from PCK and “school knowledge,” the partial correlation $r_{SC1,Ex1 > Ex2}$ turns out to be significant. This might be because we can only reveal large effects with the small sample size. However, it makes sense from a theoretical perspective. The explaining situation as it has been simulated in the used instrument was an instructional situation about physics from school curricula. We regard this as an argument for the validity of the constructed scale of joint professional knowledge that these two scales, in particular, turn out to be important, and no subscale contradicts the assumed argument that professional knowledge is a disposition for action-related skills. We conducted a post hoc analysis of the data and repeated the cross-lagged panel analysis, excluding the subdimensions “PCK-related theoretical concepts” (from PCK) and “deeper school knowledge” (from CK) (Figure 8). One could argue that the theoretical knowledge of PCK concepts such as conceptual change and a deeper understanding of school physics, its concepts,
and the limitation of its models help in other situations in the profession of a science teacher such as planning a lesson. They may, however, not be essential in explaining situations. The repeated cross-lagged panel analysis with the newly constructed scale of supposedly relevant professional knowledge just leads to a stronger claim that all four criteria used in this study are fulfilled. By including all parts of the measured professional knowledge, we may have even underestimated the relationship between professional knowledge and the development of explaining skills.

Our analysis, therefore, is rather cautious. At least, the post hoc analysis supports the assumption that the reported effects for the joint professional knowledge scale have not just been identified due to the way the joint scale for professional knowledge was constructed. A limitation of this analysis is that the sample size does not allow controlling the intercorrelation of the analyzed subscales.

One reason for these results of a relationship between professional knowledge before field experience and the development of explaining skills during field experience could, again, involve the process of reflecting about teaching (see above). Professional knowledge (particularly knowledge about instructional situations and physics in school textbooks) might provide the cognitive categories necessary to reflect on teaching productively, aiming for better instruction. In terms of the RCM, that would mean that pPCK provides the knowledge to reflect in such a way that the ePCK increases. In terms of the MoC, that would mean that the dispositions provide the categories to successfully reflect on performance—a relationship that is explicitly mentioned by Blömeke et al. (2015).

We would like to highlight that our description of how pPCK affects an increase of ePCK by reflecting about teaching is just a hypothesis. This should be researched in further studies. The main result of our study is that it depends on the level of professional knowledge how much preservice teachers benefit from field experience concerning the development of explaining skills. There is a correlation between professional knowledge before the field experience and the development of explaining skills. Also, the correlation between professional knowledge and explaining skills at one point of a time is rather low. We would argue that these results also contribute to the general discussion of the theory-practice gap. Together, they would support the assumption that explaining skills are developed from professional knowledge (thus the cross-lagged correlation) but are also skills that exist apart from professional knowledge (thus the low correlations at one point of a time). This would support the claim that—at least for explaining skills—theoretical knowledge is transformed into practical skills. For the case of explaining skills, we would, therefore, agree with Zeichner et al. (2012): field experience might foster the transformation of knowledge achieved at the university into this particular skill necessary for the profession of a teacher. However, further studies in a broader context are needed. The present paper is mainly focused on how professional knowledge affects the development of explaining skills during a field experience.

This result might also have an impact on how field experiences should be implemented in teacher education. Based on our results, we would recommend implementing field experiences to develop practical skills after proper instruction about professional knowledge—and that means in a late stage of teacher education. Of course, it remains an open question exactly how much professional knowledge is required. We would like to point out that we analyzed a field experience as a part of a master’s program and still could not identify a general increase in explaining skills. Maybe the first phase of German teacher education should not aim to develop action-related skills at all and should focus on professional knowledge instead. Field experiences with different goals, for example, to identify whether or not the profession of a physics
teacher is the choice for a lifetime, of course, could still have their place at the beginning of teacher education.

Also, based on our results, we cannot recommend teacher education models in which teachers start teaching right from the beginning without a certain instruction on CK, PCK, and PK. Professional knowledge has been identified as an important source of action-related skills, at least for explaining situations, and practical experience just helps to develop these skills after a certain level of CK, PCK, and PK has been achieved.

We would like to stress again that we only investigated the relationship between professional knowledge and one particular practical skill, explaining physics. The focus on explaining skills is a limitation of the study by design. The analysis of further data on action-related skills in reflecting and planning physics teaching might reveal more insights into the relationship between professional knowledge and action-related skills for additional fields of a physics teacher’s everyday duties. We encourage repeating the study with a different sample and/or focus on different action-related skills.

Still, the present study is, to our knowledge, the first to apply a design that allows to investigate whether or not professional knowledge (or pPCK) is necessary for action-related skills (or ePCK) in science. To estimate how important professional knowledge is as a source compared to other possible variables, a cross-lagged panel does not contribute much—experimental designs would be required. However, the fact that this assumed relationship has been identified for one practical situation supports the concept of professional knowledge as the knowledge that helps science teachers to deal with their everyday problems in general.

ORCID

Christoph Kulgemeyer https://orcid.org/0000-0001-6659-8170
Andreas Borowski https://orcid.org/0000-0002-9502-0420

REFERENCES

Abell, S. (2007). Research on science teachers’ knowledge. In S. Abell & N. Lederman(Eds.), Handbook of research on science education (pp. 1105–1149). Mahwah, NJ: Lawrence Erlbaum Associates.
Alonzo, A., Berry, A., & Nilsson, P. (2019, 2019). Unpacking the complexity of science teachers’ PCK in action: Enacted and personal PCK. In A. Hume, R. Cooper, & A. Borowski(Eds.), Repositioning pedagogical content knowledge in teachers’ knowledge for teaching science (pp. 93–116). Singapore: Springer.
Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995). Cognitive tutors: Lessons learned. The Journal of the Learning Sciences, 4, 67–207.
Aufschnaiter, C. v., & Blömeke, S. (2010). Professionelle kompetenz von (angehenden) lehrkräften erfassen—Desiderata [Measuring teachers’ professional competence—Desiderata]. Zeitschrift für Didaktik der Naturwissenschaften, 16, 361–367.
Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., ... Tsai, Y. M. (2010). Teachers’ mathematical knowledge, cognitive activation in the classroom, and student progress. American Educational Research Journal, 47(1), 133–180.
Berland, L., & McNeill, K. (2012). For whom is argument and explanation a necessary distinction? A response to Osborne and Patterson. Science Education, 96(5), 808–813.
Berthold, K., Röder, H., Knörzer, D., Kessler, W., & Renkl, A. (2011). The double-edged effects of explanation prompts. Computers in Human Behavior, 27, 69–75.
Besa, S., & Bödcher, M. (2014). Empirical evidence on field experiences in teacher education. In K.-H. Arnold, A. Gröschner, & T. Hascher(Eds.), Schulpraktika in der lehrerbildung: Theoretische konzeptionen, grundlagen und efekte [Field experiences in teacher education: Theoretical concepts, basics, and effects] (pp. 129–145). Münster, Germany: Waxmann.
Black, J. B., Carroll, J. M., & McGuigan, S. M. (1987). What kind of minimal instruction is the best? In J. M. Carroll & P. P. Tanner (Eds.), Proceedings of CHI+GI’87 Human Factors in Computing Systems and Graphics Interface (pp. 159–162). New York, NY: ACM.

Blömeke, S., Gustafsson, J., & Shavelson, R. (2015). Beyond dichotomies. Competence viewed as a continuum. Zeitschrift für Psychologie, 223(1), 3–13.

Carlson, J., & Daehler, K. (2019, 2019). The refined consensus model of pedagogical content knowledge in science education. In A. Hume, R. Cooper, & A. Borowski (Eds.), Repositioning pedagogical content knowledge in teachers’ knowledge for teaching science (pp. 93–116). 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.

Champagne, A. B., Klopfner, L. E., & Gunstone, R. F. (1982). Cognitive research and the design of science instruction. Educational Psychologist, 17, 31–53.

Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. Cognitive Science, 13, 145–182.

Chi, M. T. H., de Leeuw, N., Chiu, M. H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. Cognitive Science, 18, 439–477.

Chi, M. T. H., Siler, S. A., Jeong, H., Yamauchi, T., & Hausmann, R. G. (2001). Learning from human tutoring. Cognitive Science, 25, 471–533.

Clarke, D., & Hollingsworth, H. (2002). Elaborating a model of teacher professional growth. Teaching and Teacher Education, 18, 947–967.

Cohen, E., Hoz, R., & Kaplan, H. (2013). The practicum in preservice teacher education: A review of empirical studies. Teaching Education, 24(4), 345–380.

Delaney, S. (2012). A validation study of the use of mathematical knowledge for teaching measures in Ireland. ZDM Mathematics Education, 44(3), 427–441.

Diakidoy, I. N., Kendeou, P., & Ioannides, C. (2003). Reading about energy: The effects of text structure in science learning and conceptual change. Contemporary Educational Psychology, 28, 335–356.

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.

Ergönenc, J., Neumann, K., & Fischer, H. (2014). The impact of pedagogical content knowledge on cognitive activation and student learning. In H. E. Fischer, P. Labudde, K. Neumann, & J. Viiri (Eds.), Quality of instruction in physics. Comparing Finland, Germany and Switzerland (pp. 145–160). Waxmann: Münster, Germany.

Fischer, H., Borowski, A., & Tepper, O. (2012). Professional knowledge of science teachers. In B. Fraser, K. Tobin, & C. McRobbie (Eds.), Second international handbook of science education (pp. 435–448). Dordrecht, The Netherlands: Springer.

Fischer, H., Neumann, K., Labudde, P., & Viiri, J. (2014). Quality of instruction in physics. Comparing Finland, Germany and Switzerland. Münster, Germany: Waxmann.

Gage, N. (1968). The microcriterion of effectiveness in explaining. In N. Gage (Ed.), Explorations of the teacher’s effectiveness in explaining, technical report no. 4 (pp. 1–8). Stanford Center for Research and Development in Teaching; Stanford, CA.

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

Gramzow, Y., Riese, J., & Reinhold, P. (2013). Modellierung fachdidaktischen wissens angehender physiklehrkräfte [Modelling prospective teachers’ knowledge of physics education]. Zeitschrift für Didaktik der Naturwissenschaften, 19(1), 7–30.

Ganger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. Econometrica, 37, 424–438.

Helmke, A. (2006). Unterrichtsqualität: Erfassen, bewerten, verbessern [Teaching quality: Measurement and Improvement]. Seelze, Germany: Kallmeyersche Verlagsbuchhandlung.

Hempel, C., & Oppenheim, P. (1948). Studies in the logic of explanation. Philosophy of Science, 15(2), 135–175.
Hill, H. C., Blunk, M. L., Charalambous, C. Y., Lewis, J. M., Phelps, G. C., Sleep, L., & Ball, D. L. (2008). Mathematical knowledge for teaching and the mathematical quality of instruction: An exploratory study. *Cognition and Instruction, 26*(4), 430–511.

Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers’ mathematical knowledge for teaching on student achievement. *American Educational Research Journal, 42*(2), 371–406.

Hoffman, J. V., Wetzels, M. M., Maloch, B., Greer, E., Taylor, L., DeJulio, S., & Vlach, S. K. (2015). What can we learn from studying the coaching interactions between cooperating teachers and preservice teachers? A literature review. *Teaching and Teacher Education, 52*, 99–112.

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

Hume, A., Cooper, R., & Borowski, A.(Eds.). (2019). *Repositioning pedagogical content knowledge in teachers’ knowledge for teaching science*. Singapore: Springer.

Keller, M., Neumann, K., & Fischer, H. (2016). The impact of teachers’ pedagogical content knowledge and motivation on students’ achievement and interest. *Journal of Research in Science Teaching, 54*(5), 568–614.

Kulgemeyer, C. (2018). Towards a framework for effective instructional explanations in science teaching. *Studies in Science Education, 54*(2), 109–139. https://doi.org/10.1080/03057267.2018.1598054

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., & Schecker, H. (2009). Kommunikationskompetenz in der physik: Zur entwicklung eines domänen spezifischen kompetenzbegriffs [Science communication competence: Developing a domain-specific concept]. *Zeitschrift für Didaktik der Naturwissenschaften, 15*, 131–153.

Kulgemeyer, C., & Schecker, H. (2012). Physikalische kommunikationskompetenz—Empirische validierung eines normativen modells [Physics communication competence—Empirical validation of a normative model]. *Zeitschrift für Didaktik der Naturwissenschaften, 18*, 29–54.

Kulgemeyer, C., & Schecker, H. (2013). Students explaining science—Assessment of science communication competence. *Research in Science Education, 43*, 2235–2256.

Kulgemeyer, C., & Starauschek, E. (2014, 2019). Analyse der verständlichkeit naturwissenschaftlicher fachtexte [Analysing the comprehensibility of scientific texts]. In D. Krüger, I. Parchmann, & H. Schecker(Eds.), *Methoden in der naturwissenschaftsdidaktischen Forschung* (pp. 241–253). Heidelberg, Germany: Springer.

Kulgemeyer, C., & Tomczyszyn, E. (2015). Physik erklären—Messung der erklärens fähigkeit 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, 41*(7), 890–910.

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, The Netherlands: Kluwer Academic Publishers.

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

Nunnally, J. C. (1967). *Psychometric theory*. New York, NY: McGraw-Hill.

Oeberst, A., Haberstroh, S., & Gnams, T. (2015). Not really the same: Computerized and real lotteries in decision making research. *Computers in Human Behavior, 44*, 250–257.

Osborne, J. F., & Patterson, A. (2011). Scientific argument and explanation: A necessary distinction? *Science Education, 95*(4), 627–638.

Park, S., & Oliver, J. S. (2008). Revisiting the conceptualisation of pedagogical content knowledge (PCK): PCK as a conceptual tool to understand teachers as professionals. *Research in Science Education, 38*(3), 261–284.
Peterson, P., Carpenter, T., & Fennema, E. (1989). Teachers’ knowledge of students’ knowledge in mathematics problem solving: Correlating and case analysis. *Journal of Educational Psychology, 81*(4), 558–569.

Reinders, H. (2006). Kausalanalysen in der längsschnittforschung [Causal analysis in longitudinal studies]. *Diskurs Kindheits und Jugendforschung, 4*, 569–587.

Renkl, A. (2002). Worked-out examples: Instructional explanations support learning by self-explanations. *Learning and Instruction, 12*, 529–556.

Retelsdorf, J., Butler, R., Streblow, L., & Schiefele, U. (2010). Teachers’ goal orientations for teaching: Association with instructional practices, interest in teaching, and burnout. *Learning and Instruction, 20*, 30–46.

Riese, J., Kulgemeyer, C., Borowski, A., Fischer, H., Schecker, H., Reinhold, P., ... Zander, S. (2015). Modellierung und messung des professionswissens in der lehramtsausbildung physik [Modeling and measuring professional knowledge in physics teacher training]. *Beilfeht der Zeitschrift für Pädagogik, 61*, 55–79.

Riese, J., & Reinhold, P. (2012). Die professionelle kompetenz angehender physiklehrkräfte in verschiedenen ausbildungsförmern [The professional competencies of trainee teachers in physics in different educational programs]. *Zeitschrift für Erziehungswissenschaft, 15*, 111–143.

Roelle, J., Berthold, K., & Renkl, A. (2014). Two instructional aids to optimise processing and learning from instructional explanations. *Instructional Science, 42*, 207–228.

Sánchez, E., García Rodicio, H., & Acuña, S. R. (2009). Are instructional explanations more effective in the context of an impasse? *Instructional Science, 37*, 537–563.

Scott, P., Asoko, H., & Driver, R. (1992). Teaching for conceptual change—A review of strategies. In R. Duit, F. Goldberg, & H. Niedderer (Eds.), *Research in physics learning: Theoretical issues and empirical studies* (pp. 310–329). Kiel, Germany: IPN.

Seidel, T., & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results. *Review of Educational Research, 77*(4), 454–499.

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* (pp. 183–214). Münster, Germany: Waxmann.

Sevian, H., & Gonsalves, L. (2008). Analysing how scientists explain their research: A rubric for measuring the effectiveness of scientific explanations. *International Journal of Science Education, 30*(11), 1441–1467.

Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Education Review, 57*(1), 1–22.

Staub, F. C., & Stern, E. (2002). The nature of teachers’ pedagogical content belief matters for students’ achievement gains: Quasi-experimental evidence from elementary mathematics. *Journal of Educational Psychology, 94*, 344–355.

Tabachnick, B., & Zeichner, K. (1984). The impact of the student teaching experience on the development of teacher perspectives. *Journal of Teacher Education, 35*, 28–36.

Terhart, E. (2012). Wie wirkt Lehrerbildung? Forschungsprobleme und gestaltungsfragen [What is the impact of teacher education?]. *Zeitschrift für Bildungsforschung, 2*(1), 3–21.

Tomlinson, P. D., & Hunt, D. E. (1971). Differential effects of rule-example order as a function of learner conceptual level. *Canadian Journal of Behavioural Science, 3*(3), 237–245.

Toulmin, S. (1958). *The uses of argument*. Cambridge, England: Cambridge University Press.

Treffert, D., & Harrison, A. (1999). The genesis of effective science explanations for the classroom. In J. Lougahan (Ed.), *Researching teaching: methodologies and practices for understanding pedagogy*, (pp.28–43). Abingdon, UK: Routledge.

van Driel, J. H., Verloop, N., & de Vos, W. (1998). Developing science teachers’ pedagogical content knowledge. *Journal of Research in Science Teaching, 35*(6), 673–695.

Vogelsang, C. (2014). *Validierung eines instruments zur erfassung der professionellen handlungskompetenz von physiklehrkräften—zusammenhangsanalysen zwischen lehrerkompetenz und lehrerperformanz* [Validating an Instrument to measure physics teachers’ professional competence]. Berlin, Germany: Logos.

Vogelsang, C., Borowski, A., Kulgemeyer, C., Reinhold, P., Riese, J., & Schecker, H. (2019). Entwicklung von professionswissen und unterrichtsperformanz im lehramtsstudium physik—Analysen zu valider testwerteninterpretation [Development of professional knowledge and performance in academic physics teacher education]. *Zeitschrift für Pädagogik, 64*(4), 473–491.
Webb, N. M., Ing, M., Kersting, N., & Nemer, K. M. (2006). Help seeking in cooperative learning groups. In S. A. Karabenick & R. S. Newman (Eds.), Help seeking in academic settings: Goals, groups, and contexts (pp. 45–88). Mahwah, NJ: Lawrence Erlbaum Associates.

Webb, N. M., Troper, J. D., & Fall, R. (1995). Constructive activity and learning in collaborative small groups. *Journal of Educational Psychology, 87*, 406–423.

Weinert, F. E. (2001). Concept of competence: A conceptual clarification. In D. S. Rychen & L. H. Salganik (Eds.), *Defining and selecting key competencies* (pp. 45–66). Göttingen, Germany: Hogrefe.

Wilson, S., & Floden, R. (2003). *Creating effective teachers: Concise answers for hard questions*. New York, NY: AACTE Publications.

Wittwer, J., & Ihme, N. (2014). Reading skill moderates the impact of semantic similarity and causal specificity on the coherence of explanations. *Discourse Processes, 51*, 143–166.

Wittwer, J., & Renkl, A. (2008). Why instructional explanations often do not work: A framework for understanding the effectiveness of instructional explanations. *Educational Psychologist, 43*(1), 49–64.

Wolf, E., Harrington, K., Clark, S., & Miller, M. (2013). Sample size requirements for structural equation models: An evaluation of power, bias, and solution propriety. *Educational and Psychological Measurement, 6*, 913–934.

Zeichner, K., Payne, K., & Brayko, K. (2012). *Democratizing knowledge in university teacher education through practice-based methods teaching and mediated field experience in schools and communities*. Seattle, WA: University of Washington.

**How to cite this article:** Kulgemeyer C, Borowski A, Buschhüter D, et al. Professional knowledge affects action-related skills: The development of preservice physics teachers’ explaining skills during a field experience. *J Res Sci Teach.* 2020;1–29. [https://doi.org/10.1002/tea.21632](https://doi.org/10.1002/tea.21632)